KU LEUVEN

ARENBERG DOCTORAL SCHOOL
FACULTY OF ENGINEERING SCIENCE

Conceptual model structure identification and calibration for river and sewer systems

Vincent Wolfs

Supervisor:
Prof. dr. ir. Patrick Willems

Dissertation presented in partial fulfilment of the requirements for the degree of PhD in Engineering Science

January 2016

Cover image:

Adopted from "Water" by Joee Yabuki, licensed under CC BY-NC-SA 2.0. Original: https://www.flickr.com/photos/webcreators_info/8605953552/

CONCEPTUAL MODEL STRUCTURE IDENTIFICATION AND CALIBRATION FOR RIVER AND SEWER SYSTEMS

Vincent WOLFS

Promotor:
Prof. dr. ir. P. Willems

Members of the Examination Committee:
Prof. dr. ir. D. Vandermeulen, chair
Prof. dr. ir. I. Smets
Em. prof. dr. ir. J. Berlamont
Prof. dr. ir. E. Deleersnijder
(Université Catholique de Louvain and
Delft University of Technology)
Prof. dr. ir. P. Vanrolleghem
(Université Laval)

Dissertation presented in partial fulfilment of the requirements for the degree of PhD in Engineering Science

January 2016

Uitgegeven in eigen beheer, Vincent Wolfs, Kasteelpark Arenberg 40 – bus 2448, B-3001 Heverlee (Belgium)

Make things as simple as possible, but not simpler.

Einstein

Voorwoord

Nu mijn doctoraat bijna ten einde is, neem ik graag even de tijd om alle mensen te bedanken die, bewust of onbewust, hun steentje hebben bijgedragen aan dit werk. Zonder hen was u deze tekst niet aan het lezen.

In de eerste plaats gaat mijn dank uit naar mijn promotor Professor Patrick Willems. Tijdens het uitwerken van mijn masterproef onder uw begeleiding leerde ik dat "conceptuele modelstructuuridentificatie en -kalibratie" veel boeiender en minder theoretisch is dan de titel doet uitschijnen. Na mijn thesis bleven er echter nog veel uitdagingen over. Ik ging graag in op uw vraag of ik niet geïnteresseerd was in het uitwerken van een doctoraat rond hetzelfde onderwerp. Ik waardeer enorm de vrijheid die u me gegeven hebt om een invulling te geven aan mijn onderzoek. U maakte altijd tijd vrij om me met raad en daad bij te staan, of het nu ging over het verbeteren van teksten, het uitwerken van mijn doctoraatsplanning of, zoals u zelf vaak zei, het "fine tunen" van mijn onderzoek. Kortom, ontzettend bedankt voor de hulp, het vertrouwen en alle geboden kansen!

Ik wil Professor Jean Berlamont en Professor Ilse Smets bedanken voor het opvolgen van mijn onderzoek en het geven van suggesties van bij de start van mijn doctoraat tot en met de finale. Daarnaast wil ik ook de andere leden van mijn jury bedanken voor het geven van feedback en het nalezen van de tekst. Jullie opmerkingen hebben zeker tot een verbetering van de tekst geleid. Professor Eric Deleersnijder, hartelijk dank voor het delen van uw expertise, in het bijzonder in de domeinen waar ik minder in thuis ben. Professor Peter Vanrolleghem, bedankt voor het maken van de verre reis om mijn preliminaire verdediging bij te wonen en het stellen van kritische vragen die tot verder nadenken aanzetten. Uw enthousiasme voor onderzoek werkt inspirerend. Mijn dank gaat ook naar Professor Dirk Vandermeulen voor het voorzitten van mijn jury. Tot slot wil ik Anita bedanken voor de administratieve ondersteuning.

Ik ben het IWT erkentelijk voor de financiering van mijn onderzoek als SBO doctoraatsbursaal. Ik wil ook de bedrijven DHI en Innovyze bedanken voor het ter beschikking stellen van de licenties voor de MIKE en InfoWorks software, en de Vlaamse Milieumaatschappij, het Waterbouwkundig Laboratorium, Aquafin en Farys voor het aanleveren van modellen en data. Zonder deze modellen had ik dit onderzoek niet kunnen uitvoeren.

De voorbije jaren ben ik nooit met tegenzin naar het Kasteelpark gegaan. De collega's spelen daar zeker een belangrijke rol in. Bedankt Bert, Evert, Niels, Pieter en Rutger voor de leuke middagen en de soms al te fanatieke kubb-wedstrijdjes in weer en wind. Ik waag me niet aan het opsommen van alle andere collega's en ondertussen ook oud-collega's die ik nog wil bedanken, want ik wil niemand vergeten en dit boek is bovendien al lijvig genoeg. Bedankt allemaal voor de gezellige werksfeer, lekkere en vaak exotische traktaties bij verjaardagen, en alle fijne momenten!

Een speciaal woordje van dank gaat naar mijn ouders en zus Celine om er altijd te zijn voor mij. Jullie hebben ervoor gezorgd dat ik hier sta. Mijn toekomstige schoonouders, Marie-Jeanne en Benny, bedankt voor jullie steun. Jullie deur stond altijd open voor mij. Dankzij jullie allemaal heb ik de afgelopen vier jaar onbezorgd aan mijn doctoraat kunnen werken.

Ik wil ook mijn vrienden, oud-kotgenoten en iedereen die mijn doctoraat van nabij of wat verderaf volgde bedanken. Ben, David, Dries, Elwin, Ine, Kevin, Lies, Machteld, Nele, Peter en Steven verdienen een speciale vermelding. Jullie hebben, de ene al sinds wat langer dan de andere, gezorgd voor de nodige afwisseling.

Tot slot wil ik nog een heel bijzonder iemand bedanken. I've saved the best for last. Astrid, ik vond het moeilijk om te schrijven wat je voor mij betekent. Dankzij jouw steun, geduld en liefde heb ik – hebben we! – dit bereikt. Je vrolijkheid, optimisme en stralende lach deden de mindere momenten al snel vergeten. Ik kijk enthousiast en vol vertrouwen uit naar onze toekomst samen!

Vincent
Leuven, januari 2016

Abstract

Water management is constantly evolving. Trends, such as population growth, urbanization and climate change, pose new challenges to water management. Risks and uncertainties have to be taken into account in order to develop strategies that remain cost-effective on the long term. Water systems consist of different subsystems and interactions, which demand for an integrated approach to achieve optimal solutions. Due to the intricate complexity of developing adequate strategies, the use of mathematical models that accurately emulate reality is indispensable. Conventionally used full hydrodynamic models suffer from several fundamental shortcomings, of which a prolonged calculation time, the high model complexity and limited flexibility in terms of model adaptability and interfacing are arguably the main ones. These deficiencies compel the use of parsimonious conceptual models for numerous water management problems, which reason from simplified relationships to mimic the response of the system. These conceptual models are therefore computationally very efficient and easily adaptable to practical needs. This dissertation focuses on the development and application of such a conceptual modelling approach for river and sewer hydraulic simulations.

The first part of the dissertation presents the newly developed conceptual water quantity modelling approach, which aims to be suitable for a wide range of scenario investigations in support of water engineering and management. First, a variety of model structures was investigated, applied and improved while elaborating several case studies. Besides the in hydrology more traditionally employed model structures such as linear reservoir theory and different linear structures, special attention was paid to the performance of machine learning techniques and expert systems. Adaptive Neuro Fuzzy Inference Systems were used for lumped floodplain modelling. Artificial Neural Networks (ANNs) have proved successful in emulating complex flows in rivers, floodplains and urban drainage systems. In addition, the performance of ANNs for capturing non-

univocal behaviour in rating curve modelling was compared to that of M5' model trees and of piecewise linear relationships with state dependent parameters. Subsequently, these model structures were combined in a generic modelling framework. Two conceptual modelling approaches were developed, one for rivers and one for urban drainage systems. These approaches aim at mimicking simulation results of full hydrodynamic river and sewer models, but at a fraction of the calculation time of the detailed models. Both conceptual modelling approaches share the same fundamental modelling principles and are based on the storage cell concept, but their model topology, structures and conceptualization procedure are tailored to each system. Their modular setting, which allows the user to combine diverse incorporated model structures freely in a single model, results in powerful emulation capabilities. Complex dynamics, such as backwater effects, varying water surface profiles in rivers, and reverse and pressurized flows in sewers, can be mimicked accurately. The approaches have important mechanistic features, such as the closing of the water balance and the possible explicit incorporation of dike levels and moveable hydraulic structures of rivers. In addition, an innovative calculation scheme with an adaptive time step was developed that avoids instabilities and minimizes calculation time.

The developed approaches were integrated in a new semi-automatic software platform to configure the conceptual models, named Conceptual Model Developer (CMD). The tool guides the user via graphical user interfaces in a stepwise manner through the configuration procedure and assists the user in setting up the conceptual model topology. A close integration was foreseen with hydrodynamic software packages to import simulation results and important model characteristics. All configured model structures are assembled fully automatically in a single highly efficient calculation script written in C programming language.

The approach and software were tested extensively on several case studies in Belgium. Three of these studies are elaborated in this dissertation. The first case study quantified the impact of installing additional retention basins and varying operation controls of moveable weirs on flood probabilities along the Marke and Dender Rivers. Conceptual models were built of both rivers and consequently used for performing long term simulations, thereby accounting for antecedent conditions. The second case study used the conceptual model of the Dender River to generate flood probability maps. The approach can be incorporated in real-time flood forecasting systems. The third study analyzed the impact of

combined sewer overflows (CSOs) on the receiving river water quality in the Molse Nete catchment. Conceptual quantity models were created of the urban drainage systems of the cities of Mol and Geel, and the Molse Nete River. Next, these models were integrated and combined with conceptual water quality models for performing long term simulations. These conceptual water quality models were created in the scope of another PhD research. Other researchers also used the in this dissertation developed conceptual modelling approach for applications in which flexible simulation models with a very limited calculation time play a pivotal role. Conceptual models were configured to assess the impact on urban and river floods for three different storage strategies for the city of Turnhout. Also, model-based optimization of hydraulic structures to minimize the flood damage in the Demer catchment was elaborated using conceptual models.

Beknopte samenvatting

Waterbeheer verandert voortdurend. Trends, zoals de bevolkingsgroei, verstedelijking en de klimaatverandering, resulteren in zowel nieuwe uitdagingen als opportuniteiten voor het beheer van watersystemen. Om strategieën te ontwikkelen die ook op lange termijn effectief blijven, is het nodig om risico's en onzekerheden te kwantificeren en deze in rekening te brengen bij nieuwe ontwerpen. Gelet op de vele deelcomponenten en ruime invloedssfeer van watersystemen is een integrale aanpak noodzakelijk om optimale oplossingen te verkrijgen. Door de intrinsieke complexiteit van het ontwikkelen van strategieën is het gebruik van wiskundige modellen die de realiteit accuraat simuleren onontbeerlijk. De conventionele hydrodynamische modellen vertonen verschillende fundamentele tekortkomingen, waarvan een zeer lange rekentijd, de hoge modelcomplexiteit en gelimiteerde flexibiliteit op vlak van modelaanpasbaarheid en integratiemogelijkheden de voornaamste zijn. Deze gebreken verplichten de modelleerder om spaarse en vereenvoudigde modellen te gebruiken voor het onderzoeken en formuleren van oplossingen voor tal van problemen omtrent waterbeheer. Deze zogenaamde conceptuele modellen baseren zich op vereenvoudigde relaties om de respons van het watersysteem te beschrijven. Hierdoor zijn de conceptuele modellen bijzonder rekenefficiënt en bovendien eenvoudig aan te passen. Dit doctoraatsonderzoek richt zich op het ontwikkelen en toepassen van een conceptuele modelleringsmethodologie voor het hydraulisch simuleren van rivier- en rioleringssystemen.

Het eerste deel van het proefschrift beschrijft de nieuw ontwikkelde conceptuele modelleringsaanpak voor waterkwantiteitssimulaties. De methode beoogt inzetbaar te zijn voor een breed scala aan toepassingen ter ondersteuning van het analyseren en beheren van watersystemen. Eerst werden diverse modelstructuren onderzocht, toegepast en verbeterd via het uitvoeren van enkele gevalstudies. Naast de modelstructuren die traditioneel toegepast worden binnen de hydrologie, zoals die van de lineaire reservoirtheorie en diverse andere lineaire

structuren, werd bijzondere aandacht besteed aan technieken rond machinaal leren (*machine learning*) en expertsystemen. Zo werden Adaptieve Neuro Fuzzy Inferentie Systemen (ANFIS) gebruikt voor het modelleren van overstromingsgebieden op een ruimtelijk geaggregeerde schaal. Artificiële Neurale Netwerken (ANN) bleken geschikt voor het emuleren van meer complexe stromingen in rivieren, overstromingsgebieden en rioleringssystemen. De performantie van ANN, van stuksgewijs lineaire relaties met tijdsvariërende parameters (*state dependent parameters*) en van M5' modelbomen werd vergeleken voor het modelleren van niet-eenduidig gedrag in debiet-waterhoogte relaties. In een tweede fase werden deze modelstructuren gecombineerd tot een generieke modelleringsaanpak. De modellen die voortvloeien uit deze aanpak trachten de simulatieresultaten van gedetailleerde volledig hydrodynamische modellen na te bootsen, maar vergen slechts een fractie van de rekentijd van de hydrodynamische modellen. Twee conceptuele modelleringsmethodologieën werden ontwikkeld: één voor rivieren en één voor rioleringssystemen. Beide methodes delen dezelfde fundamentele modelleringsprincipes, maar de conceptualisering, modeltopologie en –structuren werden specifiek afgestemd op ieder systeem. Modelleerders kunnen de diverse modelstructuren vrij combineren tot een conceptueel model. Deze modulaire aanpak is bijzonder flexibel en krachtig. Complexere dynamica, zoals *backwater* effecten, variërende verhanglijnen in rivieren en stroming onder druk in rioleringsleidingen kunnen nauwkeurig gesimuleerd worden. De ontwikkelde methodes hebben belangrijke mechanistische kenmerken, zoals het sluiten van de waterbalans en de mogelijkheid om dijkhoogtes en controleerbare hydraulische structuren in rivieren expliciet te integreren in de conceptuele modellen. Daarnaast werd een innovatief rekenschema ontwikkeld dat instabiliteiten vermijdt tijdens simulaties en streeft naar een minimale rekentijd door het gebruik van een adaptieve tijdstap.

De ontwikkelde modelleringsmethodologieën werden geïntegreerd in een nieuw ontwikkeld semiautomatisch software platform voor het configureren van de conceptuele modellen, genaamd *Conceptual Model Developer* (CMD). Het programma leidt de gebruiker stapsgewijs door het configuratieproces. Er is een nauwe integratie voorzien met de hydrodynamische softwarepakketten *InfoWorks* en *MIKE* om simulatieresultaten en belangrijke modelkarakteristieken te importeren. Na identificatie en kalibratie worden de modelstructuren automatisch geassembleerd in een efficiënt rekenscript geschreven in de programmeertaal C.

De ontwikkelde aanpak en software werden uitgebreid getest op verschillende gevalstudies in België. Dit proefschrift beschrijft drie van deze studies. De eerste studie kwantificeerde de impact van het installeren van bijkomende wachtbekkens en verschillende regelobjectieven van twee controleerbare hydraulische structuren op het overstromingsrisico langs de rivieren Marke en Dender. Conceptuele modellen werden opgesteld voor beide rivieren, waarna deze gebruikt werden voor het uitvoeren van langetermijnsimulaties. Dergelijke langetermijnsimulaties maken het mogelijk om rekening te houden met antecedente systeemcondities. De tweede gevalstudie gebruikte het conceptuele model van de Dender om overstromingskanskaarten te genereren. Deze aanpak kan geïntegreerd worden in realtime overstromingswaarschuwingssystemen. De derde gevalstudie analyseerde de impact van overstorten van een gecombineerd rioleringssysteem op de waterkwaliteit van de ontvangende rivier in het stroomgebied van de Molse Nete. Hiertoe werden conceptuele hydraulische modellen opgesteld van de rioleringssystemen van de steden Mol en Geel, en van de Molse Nete rivier. Vervolgens werden deze modellen geïntegreerd en gecombineerd met een conceptueel waterkwaliteitsmodel voor het uitvoeren van een langetermijnsimulatie. Dit conceptueel waterkwaliteitsmodel werd in het kader van een ander doctoraatsonderzoek geconfigureerd.

Andere onderzoekers gebruikten de in dit doctoraatsonderzoek nieuw ontwikkelde conceptuele modelleringsaanpak reeds voor diverse toepassingen waarbij flexibele simulatiemodellen met een zeer korte rekentijd een cruciale rol speelden. Zo werden conceptuele modellen opgesteld om de impact van drie verschillende bergingsstrategieën voor de omgeving van de stad Turnhout te analyseren op stedelijke en landelijke overstromingen. Conceptuele modellen gegenereerd met de ontwikkelde methodologie werden ook gebruikt voor het modelgebaseerd optimaliseren van de regeling van controleerbare hydraulische structuren om de overstromingsschade in het Demerbekken te minimaliseren.

Acronyms

ANFIS	Adaptive Neuro Fuzzy Inference Systems
ANN	Artificial Neural Network
CFD	Computational Fluid Dynamics
CM	Conceptual Model
CMD	Conceptual Model Developer. A semi-automatic software tool to configure conceptual models.
CSO	Combined Sewer Overflow
DBM	Data-based mechanistic approach.
DEM	Digital Elevation Model
DHI	Danish Hydraulic Institute
FIS	Fuzzy Inference System
GA	Genetic Algorithm
GIS	Geographic Information System
GL	Gate Level
GP	Grid Partitioning (ANFIS)
GUI	Graphical User Interface
H	Water Level
HC	Hypsometric Curve
HS	Hydraulic Structure equations, i.e. relationships that describe the flow over a certain hydraulic structure.
IPCC	Intergovernmental Panel on Climate Change
IR	Impulse Response
IUH	Instantaneous Unit Hydrograph

IWCS/IWRS	InfoWorks CS and RS
M5'	M5' Model Tree
MF	Membership Function (ANFIS)
MPC	Model Predictive Control
NAM	Nedbør-Afstrømnings Model (in Danish)
NLT	Non-Linear Transformation
NSE	Nash-Sutcliffe Efficiency (see also Eq. 2.7)
PDM	Probability Distributed Model
PLC	Programmable Logic Controls
PWL	Piecewise Linear Relationship
Q	Discharge
RC	Rating Curve
RIV	Refined Instrumental Variable algorithm
RMSE	Root Mean Square Error
RR	Rainfall runoff
SC	Subcatchment
SClust	Subtractive Clustering (ANFIS)
SDP	State Dependent Parameter
SDP-RC	State Dependent Parameter Rating Curve
SRC	Simple Rating Curve
TF	Transfer Function
TS	Takagi Sugeno (ANFIS)
TVP	Time Varying Parameter
VHA	Vlaamse Hydrografische Atlas
VHM	Veralgemeend Hydrologisch Model (in Dutch)
VMM	Flemish Environment Agency (Vlaamse Milieu Maatschappij)
WL	Water Level
WSP	Water Surface Profile
WWTP	Waste Water Treatment Plant

Table of contents

1

Introduction

1.1 Water management

Efficient, sustainable and future-oriented water management is an important need of society. Such management aims at developing adequate strategies that can meet different objectives at the integrated water system level. Firstly, water management tries to reduce the *flood* risk and damage along rivers, in coastal zones and urban areas. Floods are probably the number one cause of losses worldwide due to natural events (Kron, 2002, 2005). The 1931 Central China Floods are an unfortunate example of the destructive power and danger of floods. With estimates ranging from 145.000 to nearly 4 million casualties (Glantz, 2003), this flood is considered the deadliest natural disaster of the 20th century (Pietz, 2002). Flanders is also regularly afflicted by floods with major economic losses. The storm of November 2010, for instance, resulted in a damage estimated at 238.146.000 US$ (EM-DAT). On the other hand, water managers aim at reducing the risk of extreme *droughts* and at ensuring *water availability* for drinking water production, domestic, industrial and agricultural use. Contrary to popular belief, Flanders is a region with very limited water availability. The average per capita availability is even scarcer than more southern European countries such as Spain, Turkey and Greece. This is mainly caused by the high population density and the absence of very large rivers in Belgium (MIRA, 2011). Another objective of water management is to achieve a good *ecological* status in all water systems as contemplated by the European Water Framework Directive (2000/60/EC) since 2000.

The rapid urban expansion and changing environmental conditions pose new challenges to water management. The area in Flanders is being built-up rapidly: the percentage of built-up land increased from 7.2% in 1976, to 11.7% in 1988 and 18.3% in 2000. Projections show that 30% to 50% of the area of Flanders will be built-up by 2050 (Poelmans et al., 2010). This sprawling trend of increasing built-up area is expected to continue in other European countries as well due to a predicted growth of the population (EEA, 2006). It is expected that more than 15% of the total territory of Western Europe will be occupied by impermeable surfaces by 2030 (WUR, MNP, 2007). Increased built-up surfaces lead to a faster runoff and consequently higher peak discharges, which in turn can lead to floods. Urban development and expansion also enhance the potential economic losses caused by floods. Secondly, climate change can have a major impact on hydrology. A worldwide increase in the frequency of extreme rain storms as a result of global warming for the late 20th century is reported, and it is expected that this trend will continue in the 21st century (IPCC, 2007 and 2012). Studies in Flanders have shown that winters will become wetter, summers drier but with more extreme rainfall events (Willems et al., 2010; Vansteenkiste 2011a,b; MIRA, 2015; Tabari et al., 2015). In addition, the expected accelerated global mean sea level rise will change coasts and can endanger coastal populations, infrastructures, ecosystems and influence rivers (e.g. McLean et al., 2001).

Practices in water management have evolved fundamentally over time. In the past, strategies were developed on a local scale and with a limited scope. Well-known examples hereof are the rectification and diking of rivers, enabling cultivation and urbanization of natural floodplains. The adverse consequences of such interventions are still noticeable. Sewer systems were purely designed for draining cities and fulfilling a sanitary function, thereby neglecting the disastrous water quality deterioration when overflows spill into rivers during heavy rainfall. Excessive groundwater extraction through pumping in phreatic and deeper confined aquifers for industry, agriculture and water production resulted in a decline of the groundwater table, a loss in pressure and salinization in coastal zones (VMM, 2006; MIRA, 2015).

Present water management strives for efficient and optimal solutions that account for the implications of designs on a larger scale and context. The need for sustainable and adaptive strategies is broadly acknowledged, which can be altered after their realization based on future changes and needs. To achieve such

management, modelling tools and a decision support system are required that comply with the following requirements:

- A **transdisciplinary and integrated system approach** is required, since the water system inherently influences numerous other (also non-aquatic) factors. Such integrated modelling not only considers the water quantity and quality in the urban drainage or river systems, but also the waste water treatment plant (WWTP), more downstream receiving water bodies and possibly even urban planning, the societal system (Wong & Brown, 2009; Ferguson et al., 2013) and economic factors (e.g. Hauger et al., 2002). Developing optimal strategies thus requires broadening the investigated scale and scope.

- The system should enable performing various **scenario analyses and optimizations**. Through the multidisciplinary approach, global cost/benefit analyses should lead to optimal strategies under varying conditions and constraints. Performing long term simulations is necessary to account for the variability in rainfall and other (antecedent) conditions in scenario analyses (e.g. Harremoës and Rauch, 1996).

- **Risks and uncertainties** should be accounted for. For instance, it is known that climate change increases temporal variations in rainfall, evapotranspiration and water availability, but the precise progress and impact of these changes remain highly uncertain. Developing adequate and cost effective management strategies requires a probabilistic framework that can account for these and other uncertainties and risks.

- The general framework and different components should be **flexible.** The system should allow easy integration and efficient use of self-defined modules and new water management techniques, such as sustainable urban drainage systems. It should be easy to expand, use and maintain the different components of the framework. Therefore, these elements and their interactions should be transparent and strike the balance between parsimony and completeness.

The importance and potential profits of such an efficient, sustainable and integrated approach are being acknowledged increasingly by both practitioners and researchers (e.g. Vlachos et al., 2001; Rauch et al., 2002; Mannina et al., 2004; Mitchell, 2004; Vanrolleghem et al., 2005; Fratini et al., 2012; Bach et al., 2014;

Dotto et al., 2014; Lawson et al., 2014; Urich and Rauch, 2014). This paradigm shift from end-of-pipe approaches and traditional technologies to more holistic and prescient designs is also reflected by changes in legislations, such as the EU Water Framework Directive (2000/60/EC) and the Floods Directive (2007/60/EC).

1.2 Water quantity modelling

Developing cost-effective, sustainable and future-proof water management strategies is very challenging due to the complexity of the system, different interactions and various objectives. Therefore, mathematical models are required that can accurately simulate the response of the system. Such models allow explaining the dynamics of a water system, identifying sources of problems and how to remedy them, assessing different strategies and analyzing the impact of changes. Mathematical models are one of the core elements in the decision making process to elaborate management strategies.

Integrated modelling, which considers different aspects of the water system itself and influenced fields (such as economical, ecological and societal aspects), is very complex and comprises various models and associated difficulties and pitfalls (e.g. Parker et al., 2002; Jakeman and Letcher, 2003; Laniak et al., 2013; Voinov and Shugart, 2013; Bach et al., 2014). Therefore, this research is confined to solely water quantity modelling of rivers and urban drainage systems. Zoppou (2001) emphasized the importance of reliable flow predictions due to strong interactions between quantity and quality processes, which was also confirmed by other studies (e.g. Rauch and Harremoës, 1996; Schütze et al., 1996; Vanrolleghem et al., 1996).

1.2.1 Model requirements

The requirements of an effective and future-oriented decision support system for water management (see §1.1) are translated into different requisite model characteristics. The models should meet the following requirements to ensure their usability and fitness in the intended decision support framework:

- simulate the flow and other state variables in the system **accurately**, including backwater and other relevant dynamic effects;

- simulate events with **very short calculation times**. Many applications in water management share the need for computationally efficient models. For instance, optimization problems and probabilistic analyses often require numerous iterations, while assessing different scenarios statistically requires long term simulations.
- have a **flexible** model structure to incorporate new technologies and to adjust the level of model detail as desired; The models should focus on the **dominant processes** to ensure parsimony, clarity and efficiency (e.g. Vanrolleghem et al., 2005; Muschalla et al., 2015), and be tailored to the intended application. Indeed, including all processes in detail is not necessary for most applications (Harremoës and Rauch, 1996; Rauch, 2002): an approach at river basin scale for instance generally requires analyses with lower model detail compared to the investigation of certain subcomponent interactions, such as between rivers and sewers.
- suitable for various **scenario investigations and optimizing strategies.** In turn, this requirement imposes demands on the incorporated model structures and parameters. For instance, optimizing the operation of controllable hydraulic structures generally requires the explicit incorporation of hydraulic structures and their parameters in the model to ensure constraints can be accounted for. Dike levels often play a crucial role in flood control strategies, and should therefore be included in the model directly.
- easy **linkable** with other models to perform integrated and multidisciplinary analyses;
- **fast and easy configuration**. The model structure should allow semi-automatic configuration.

1.2.2 Synopsis of water quantity models

A vast variety of different water quantity models exists, often with overlapping characteristics, features or objectives. Classification is required to gain insight into their features, capabilities and different modelling perspectives. Such classification is, however, not exact and different hydrologists may employ different definitions. A plethora of divergent classifications can be found in literature (e.g. Woolhiser, 1973; Fleming, 1975; Shaw, 1983; Chow, 1988; Willems, 2000; Cunderlik, 2003; Wagener et al., 2004). The aim of this section is to give a brief overview of the taxonomy of deterministic modelling often found in literature

(based on Abbott and Refsgaard, 1996). Note that many models and codes do not fit exactly in the described classification.

A suggested principal division distinguishes deterministic and stochastic models. The latter will not be discussed further in this overview and only deterministic models will be treated. These models can be partitioned according to their spatial representation of the system (distributed or lumped) and their description of relevant physical processes (empirical, conceptual and physically-based, depending on the model detail and mechanistic characteristics). Since most physically-based and conceptual models are distributed and lumped respectively, Abbott and Refsgaard (1996) propose the following division:

- Distributed physically-based models (white box)
- Lumped conceptual models (grey box)
- Empirical models (black box)

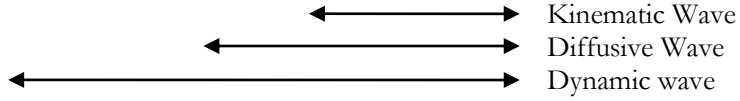
1.2.2.1 Physically-based models

The physically-based models rely on the governing continuum (partial differential) equations to calculate the flow and energy. The Navier-Stokes equations are the basic governing equations for describing motion of viscous fluids in three dimensions. The set of coupled differential equations were independently derived in the early 19th century by G.G. Stokes in England and M. Navier in France. The equations only admit a limited number of analytical solutions. Solving the full equations exactly is very difficult due to the non-linear terms. Due to their complicatedness, the Navier-Stokes equations form the subject of one of the most difficult open mathematical problems (elected as a Millenium Prize by the Clay Mathematics Institute; see also Fefferman, 2000). Therefore, the equations are generally solved numerically, which is referred to as Computational Fluid Dynamics (CFD). However, approximations of the full Navier-Stokes equations, such as the two dimensional shallow water equations and the further simplified de Saint-Venant equations (1871), yield estimates with satisfactory accuracy for most practical applications. The de Saint-Venant equations lay the theoretical foundation of nowadays channel routing and consist of the conservation of mass equation:

$$\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial z} = 0 \tag{1.1}$$

and the conservation of momentum equation:

$$\frac{\partial Q}{\partial t} + \frac{\partial}{\partial z}\frac{Q^2}{A} + gA\frac{\partial h}{\partial z} + gA\left(S_f - S_0\right) = 0 \qquad (1.2)$$



in which $Q(z,t)$ is the discharge [$m^3 \cdot s^{-1}$], t is time [s], z is the distance along the longitudinal axis of the waterway [m], $A(z,t)$ is the cross-sectional flow area [m^2], g is the gravitational acceleration [$m \cdot s^{-2}$], $h(z,t)$ is the water surface elevation above a datum or reference level [m], S_0 is the bed slope [-] and S_f is the friction slope [-]. The latter may be evaluated using a steady flow empirical formula such as the Chézy or Manning equation (see e.g. Chow, 1959; Chanson, 2004; Chaudry, 2008). The equations are valid under several assumptions (see e.g. Cunge et al., 1980; Sturm, 2001; Chaudry, 2008).

The de Saint-Venant equations are widely employed and used exclusively in hydrodynamic software, such as MIKE11, InfoWorks, HEC-RAS and SWMM. These models are referred to as "detailed" or "hydrodynamic" models in this research. As for the shallow water equations, solving the de Saint-Venant equations analytically is only feasible for some simple cases. Hence, numerical integration is inevitable. Different approaches are elaborated and used to solve the de Saint-Venant equations numerically. The MIKE11 software incorporates for instance a numerical 6-point Abbott-Ionescu scheme (Abbott and Ionescu, 1967) with self-adapting time step by monitoring the Courant number. The de Saint-Venant equations are not valid when the flow encounters sharp discontinuities, such as near hydraulic structures (e.g. Breckpot, 2013). Instead, the MIKE11 software treats such locations as internal boundaries, and hydraulic structures are preprocessed into a family of static algebraic relationships. Other software programs use similar numerical schemes.

Numerical solution of the full de Saint-Venant equations requires significant computational resources and time. Therefore, two simplifications exist that omit one or more terms of the full formulation (see Eq. 1.2) in favor of computational efficiency. The diffusive wave approximation assumes that the inertial terms are less than the gravity, friction and pressure terms, which is valid for primarily subcritical flows and slow to moderately rising flood waves in channels of rather uniform geometry (Fread, 1985). Backwater effects can still be modelled through

the $\partial h/\partial z$ –term. The kinematic wave formulation neglects additionally the hydrostatic gradient term at the expense of backwater effect calculation. Hence, this approximation assumes uniform flow with the friction slope being equal to the slope of the channel. Also, different models were developed by linearizing the de Saint-Venant equations around uniform flow conditions and by neglecting or simplifying the downstream boundary condition or terms of the continuity and momentum equations (Fread, 1985; Meunier et al., 2007), such as the models by Lighthill and Whitham (1955), Harley (1967), Keefer and McQuivey (1974), Dooge et al. (1987), Moussa (1996) and Tsai (2003). A comprehensive summary and discussion of these simplifications can be found in Miller and Yevjevich (1975) and Fread (1985). Naturally, all these simplifications invoke stringent limitations on the conditions for which their solution is valid. Therefore, models were developed that switch between different levels of simplifications depending on the routing conditions, such as the SURKNET model by Pansic and Yen (1982), the SERAIL model of Chocat et al. (1983) and a storage model by Motiee et al. (1997).

Note that the de Saint-Venant equations assume the flow to be one dimensional, but flow in large rivers and floodplains is two dimensional (e.g. Chaudry, 2008). Quasi 2-D approaches exist that describe floodplains using an additional network of streams, while lateral flow channels represent dikes (e.g. Timbe, 2007). In the present work, we assume that the solution of the de Saint-Venant equations gives a reasonably good approximation under these circumstances. This is a valid reasoning for most rivers in Flanders, which are generally controlled by hydraulic structures. Secondly, the de Saint-Venant equations do not describe the flow in pressurized pipes. To calculate the flow through pipes, a Preissmann slot is added to each pipe in most hydrodynamic calculations. The Preissmann slot allows a free surface to be maintained and obviates the transition to pressurized flow, enabling the use of the de Saint-Venant equations.

1.2.2.2 Empirical models

The empirical models (black box) are situated on the other end of the modelling spectrum. They are derived entirely from input-output series and are not directly based on physical processes. Consequently, most of the parameters lack a direct physical interpretation. Generally, these models are characterized by simple model structures. The best result is achieved for slowly varying rivers, with negligible inflows, backwater effects or floodplain-river interactions. The accuracy and extrapolation performance strongly depends on the calibration data and the

selected model structure and variables. The empirical models can roughly be divided into two main groups (Abbott and Refsgaard, 1996).

The first group comprises the statistically based approaches and mainly consists of regression and correlation methods. The simplest models relate the outflow to the inflow within a routing reach using lag models (e.g. Tatum, 1940; Harris, 1970). Autoregressive (Integrated) Moving Average (ARMA and ARIMA; Box and Jenkins, 1970 and 1976) models, Constrained Linear Systems (CLS; Todini and Wallis (1977)) and gauge-to-gauge relationships (e.g. Linsley et al., 1949) have been developed extensively for flood routing. The latter relate one or more states, such as flow and water elevations, at a downstream location in the river to states at an upstream station (e.g. Goswami and O'Connor, 2007; Archer and Fowler, 2008). These methods are often the most used modelling method for many large rivers in certain countries, such as India and Bangladesh (Abbott and Refsgaard, 1996). A plethora of more advanced methods are still being developed nowadays using traditional statistical methods, including approaches that also quantify the uncertainty (e.g. Porporato and Ridolfi, 2001; Krzysztofowicz, 2002; Romanowicz et al., 2008; Beven et al., 2009).

The second group consists of hydroinformatics based models, such as machine learning techniques and evolutionary algorithms. Such techniques are self-learning and their model structures can adapt themselves to data provided during configuration. The use of adaptive neuro fuzzy inference systems (e.g. Bazartseren et al., 2003; Chang and Chang, 2006) and neural networks in particular is increasingly popular in water resources modelling (see e.g. Maier and Dandy (2000) and Maier et al. (2010) for a comprehensive overview of neural networks used for forecasting water resources variables).

1.2.2.3 Conceptual models

The group of conceptual models finds itself between the (distributed) physically-based and empirical models, and seeks for a compromise between these two groups. The conceptual models can be regarded as a macroscopic representation of reality, meaning that they represent processes as they are observed. Hence, abstraction is made from several microscopic processes that underlie the observations themselves (Willems, 2000). Again, this category comprises a wide variety of models, but share basic principles. In contrary to the empirical models, these models emanate from the continuity equation by using mutually interconnected storages that represent physical elements in the catchment. Hence,

the water balance is closed during simulations. Instead of applying the physically-based formulation of the momentum equation (Eq. 1.2) discussed in the first group, a self-defined (possibly empirical) relationship is defined to calculate the flow using one or more states of the system. Most of the used model structures are simple but mechanistic, meaning that they and their parameters can be interpreted in physically meaningful terms. In contrast to the fully physically-based models, the vast majority of these parameters cannot be measured directly, but need calibration from field data or simulation results of more detailed models.

The simplest conceptual method is a storage routing model that relates the storage to solely the outflow. The first development of such model is attributed to Puls (1928) and Goodrich (1931). The well-known and ubiquitously used Muskingum method, first developed by McCarthy (1938), links the storage to both the inflow and outflow. This approach was later improved by Cunge (1969) and resulted in the Muskingum-Cunge approach (Miller and Cunge, 1975; Weinmann and Laurenson, 1979). Over time, numerous alternative descriptions and variations of both Muskingum methodologies were developed (e.g. Linsley et al. 1949; Nash, 1959; Diskin, 1967; Koussis, 1978 and 1980; Ponce and Yevjevich, 1978; Strupczewski and Kundzewicz, 1980; Dooge et al., 1982; Khan, 1993; Ponce and Chaganti, 1994; Birkhead and James, 2002; Achleitner et al., 2007; Kumar et al., 2011; Franchini et al., 2011). A noted variant of the Muskingum approach is the Kalinin-Milyukov approach (Miller and Cunge, 1975), developed in 1958 in the USSR (Fread, 1985). Models were also developed that delay the upstream incoming flow, denoted as lag route models (e.g. Linsley et al., 1949; Quick and Pipes, 1975). Most of these approaches are direct variants of the kinematic wave or diffusion analogy models as demonstrated by Weinmann and Laurenson (1979). Later, raster-based (often flood inundation) models were created that use Cunge-type storage cells (Cunge, 1975) in combination with simple equations to calculate the intercell fluxes (e.g. Estrela and Quintas, 1994; Bechteler et al., 1994; Hunter et al., 2007 and 2008; and references therein). The LISFLOOD-FP approach proposed by Bates and de Roo (2000) is arguably the most well-known and adopted conceptual raster-based model. Since calibration and validation of conceptual models can be time-consuming, some approaches were incorporated in different software packages, such as Flupol (Bujon et al., 1992), CityDrain 3 (Burger et al., 2010), Aquacycle (Mitchell et al., 2001), KOSIM (ITWH, 2000), KOSIM-WEST (Solvi, 2007), WaterAspects (Grum et al., 2004) and SMUSI (Muschalla et al., 2006).

All existing modelling approaches rely on a rather simple and limited set of equations for flood routing, and therefore have limited accuracy. The ability of modelling backwater effects for instance is of great importance, especially in urban drainage modelling, since they can have a significant impact on the system's dynamics. Vanrolleghem et al. (2009), among others, demonstrated that backwater effects can induce CSO events upstream, but also reduce CSO peaks due to the activation of storage volume in the sewer system. Therefore, attempts have been made to simulate backwater effects in conceptual models (e.g. the combiner-splitter module created by Solvi (2007)), but additional calibration is necessary, the usability is limited and flow routing remains highly simplified with the associated limitations (e.g. Duchesne et al., 2001; Vanrolleghem et al., 2009).

1.2.3 Past, present and future use of water quantity models

As for water management, the modelling of water systems has changed significantly over time. The de Saint-Venant equations, the basic theory for one-dimensional analysis of flood wave propagation, were already formulated in 1871. However, due to the mathematical complexity of these equations, simplifications were necessary to obtain solutions. Consequently, a profusion of simplified, mainly empirical, routing models appeared in literature (Fread, 1985). Initially, such very simplified modelling was necessary to alleviate the burden of solving systems of equations. These models were in general not physically based and lacked accuracy in most cases.

With the advent of computational resources since the 1950s and 1960s, more complex models gained interest and popularity. First, simplified models were still used because of the limited available computational power, but more detail was added to the models and different conceptual models were developed. At the same time, the principal theoretical work on the kinematic formulation of the de Saint-Venant equations was done by Lighthill and Whitham (1955). They named the theory "kinematic wave" and investigated the applicability of the approach extensively (Miller, 1984). Rather simple characteristics of flood waves could be simulated, such as the celerity and crest attenuation. More complex dynamics, such as backwater and inertia effects, were ignored. To include those phenomena in simulations, more elaborate versions and eventually the full solution of the de Saint-Venant equations are needed. Stoker (1953) and Isaacson et al. (1954) were the first to use the full de Saint-Venant formulation to simulate flood routing on a real case. Since then, finding numerical solutions of the de Saint-Venant equations has become more and more popular and efficient due to the ever

increasing computational power. Since several decades, hydrodynamic software is the standard tool for flood routing for most practitioners and researchers.

For a long time, one of the primary objectives of modellers was to include as much detail as possible in their models in pursuit of the most accurate emulation of reality. Complex hydrodynamic models were increasingly prevalent. Nowadays, simplified models gain attention again. Many advocate the use of simpler models and stress the need for pragmatic and parsimonious models (e.g. Willems and Berlamont, 2002; Wagener et al., 2004; Freni et al., 2010; Bach et al., 2014). This shift to simpler models is driven by several major disadvantages of detailed models when used in water management applications. Three principal and fundamental shortcomings of detailed models are briefly discussed.

- Hydrodynamic models suffer from **prolonged calculation times**, which is arguably the most critical problem. The long simulation times impede the use of detailed models for applications requiring a large number of iterations, or for performing long term simulations. Such applications and analyses are essential to deal with uncertainties, risks, scenario and impact investigations, and optimization questions.

- Hydrodynamic models are highly **inflexible** due to the character of the underlying equations. The level of model detail cannot be adjusted given the amount and accuracy of available data or specified objectives. Therefore, models are often **overly complex** for the intended application. The inflexible nature also results in imbalanced models: model components with ample calibration data and hence lower uncertainties are interfaced with simpler components that lack accuracy. Similarly, processes often act on different time and spatial scales. Momentary sewer overflows result in an acute burdening and deterioration of the water quality of the receiving river, which could persist for much longer times. The spatial scales in integrated urban drainage models vary from small, individual households to the design of centralized SUDS and up to the receiving river catchment scale (Makropoulos et al., 2008). The non-adaptable model structures and equations of detailed models are too rigid to efficiently deal with such varying scales and uncertainties.

- The detailed models are characterized by a rigid and closed model structure and implementation, which makes **interfacing and**

integration with other models often difficult. Due to the typically closed programming implementation of detailed models, the interfacing with other models and platforms is technically challenging or even impossible. Workarounds are being developed, such as Delft-FEWS, but the use of such platforms is often cumbersome and limited. In many cases, bilateral interactions, in which two interfaced models can influence each other at every time step during a simulation, can often not be accounted for. In addition, linking complex models leads to an even increased model complexity and calculation time.

Even if computational resources evolve significantly during the next decade, these issues, including the long calculation times, will persist and remain problematic for numerous applications in water management. The detailed models do not meet the model requirements outlined in §1.2.1. Mainly due to their flexibility and computational efficiency, simplified models overcome these limitations. The importance of simplified modelling approaches is also reflected by the large number of studies and analyses in which they play a pivotal role. Optimization problems necessitate fast quantitative predictions of the river's state, such as conjunctive water use and resource management (e.g. Peralta et al., 1995; Fredericks et al., 1998; Rejani et al., 2009; Montazar et al., 2010), operation of reservoir systems (e.g. Yeh, 1985; Faber and Stedinger, 2001; Forster et al., 2008; Alemu et al., 2011; Castelletti et al., 2012a) and operation of hydraulic structures for flood control (e.g. Breckpot et al., 2013; Van den Zegel et al., 2014; Yazdi and Salehi Neyshabouri, 2014). Efficient river models also form a vital tool to generate real time flood forecasts (e.g. Pedregal et al., 2009), to predict the river's state for different input time series such as for ensembles of rainfall forecasts (e.g. Demeritt et al., 2007; Cloke and Pappenberger, 2009; Van Steenbergen and Willems, 2014) and to asses various impact and scenario analyses, such as climate or land use change (e.g. Ashley et al., 2005; Booji, 2005; Dobler et al., 2012). Models with short calculation times are also being used in many applications in the field of urban drainage, including uncertainty assessments (Willems, 2008; Achleitner et al., 2009; Freni et al., 2009), sensitivity analyses and auto-calibration of model parameters (Kanso et al. 2003; Kleidorfer et al., 2009; Freni and Mannina, 2010), and real-time control (RTC) and other multi-objective optimization problems (Vanrolleghem et al., 2005; Borsányi et al., 2008; Fu et al., 2009; Dirckx et al., 2011; Vezzaro et al., 2014).

It is clear that each modelling approach is bound by its limitations, and thus no universally superior model exists. The choice for a model type depends on the

objective, the considered problem, the knowledge of the modeller and the available computational resources and data (Yen, 1994). The selection, implementation and interfacing of models should be carefully tailored to the intended research or application. A mixture of models, each with its own strengths, is needed to tackle the present and future challenges in water management.

1.3 Towards a generic framework for decision support for water management

A generic framework is proposed for the mathematical modelling related component of decision support systems for an efficient, sustainable and future-oriented water management. The framework builds further on the ideas proposed and applied in Willems (2000) and related studies, and is based on the complementary use of conceptual and detailed hydrodynamic models to fully benefit from the advantages of both model types. Such synergetic approach allows for the efficient development of optimal water management strategies. Figure 1.1 schematizes the proposed framework for integrated modelling in support of water management analyses and planning.

As discussed in §1.2.3, detailed hydrodynamic models are unsuited for many applications in water management. The significant shortcomings of the conventionally used detailed models impose the use of alternative and simplified modelling approaches. We advocate the use of flexible surrogate models, which emulate the physically-based detailed models, but rely on much simpler and more parsimonious relationships. Due to their uncomplicated characterization, their calculation time is several orders of magnitude shorter than that of detailed models. Additionally, simplified models make use of more flexible model structures that can be adapted easier. All these features enable these models to be integrated on large scales. To ensure these simplified models meet the requirements identified in §1.2.1 and are thus effectively usable for developing future-oriented and optimal water management strategies, only conceptual-type models can be employed. In contrary to the even simpler empirical models, conceptual models still have important and requisite mechanistic features. Hence, their main characteristics and advantages make conceptual models well suited to be used in decision support systems and other applications requiring fast (integrated) models.

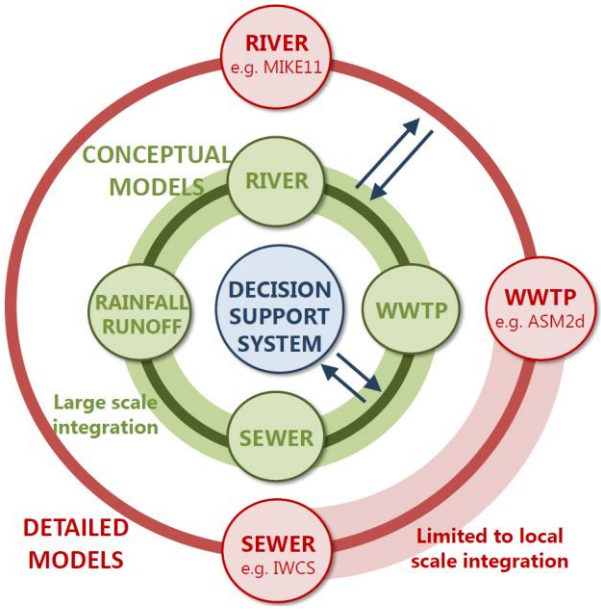


Figure 1.1: Generic framework based on the complementary use of conceptual and detailed models to develop water management strategies. (adopted from Willems, 2000)

Conceptual-type models cannot replace detailed models entirely though. Firstly, the configuration of conceptual models demands a significant amount of data because many parameters lack a direct physical interpretation (e.g. Willems, 2006; Hunter et al., 2007). Collecting all requisite data through measurement campaigns is an expensive and time-consuming task (Vanrolleghem et al., 1999), and consequently infeasible in nearly all cases. Therefore, using simulation results of hydrodynamic models to configure conceptual-type models has become standard practice to compensate for the lack of accurate measurement data (e.g. Vaes, 1999; Meirlaen et al., 2001). Secondly, physically-based models remain necessary for spatially detailed analyses, since conceptual model aggregate processes and only give a lumped representation of the system. This bilateral interplay between the conceptual and detailed models is illustrated by Figure 1.1: conceptual models used in decision support of water management applications emulate the results of the physically-based models, while the latter can be used to assess the outcome of such analyses on finer scales or with enhanced accuracy.

Figure 1.2 schematizes the proposed conceptual modelling framework. Each component, such as the river or sewer system, is transformed into a conceptual

model. The mechanistic nature of the conceptual models must allow analyzing different scenarios and model adaptations, such as the installation of dikes and retention basins, applying a water sensitive urban design approach in cities and optimizing gate regulations of controllable hydraulic structures. Next, these individual models are interfaced at catchment level, enabling integrated analyses. In addition, the flexible framework should allow to plug-in other structures and user-defined modules, such as a waste water treatment plant or water production center. Note that the hydrological component is also indicated on the figure, although catchment hydrology is not investigated in this research. Instead, different existing hydrological models are used throughout the research.

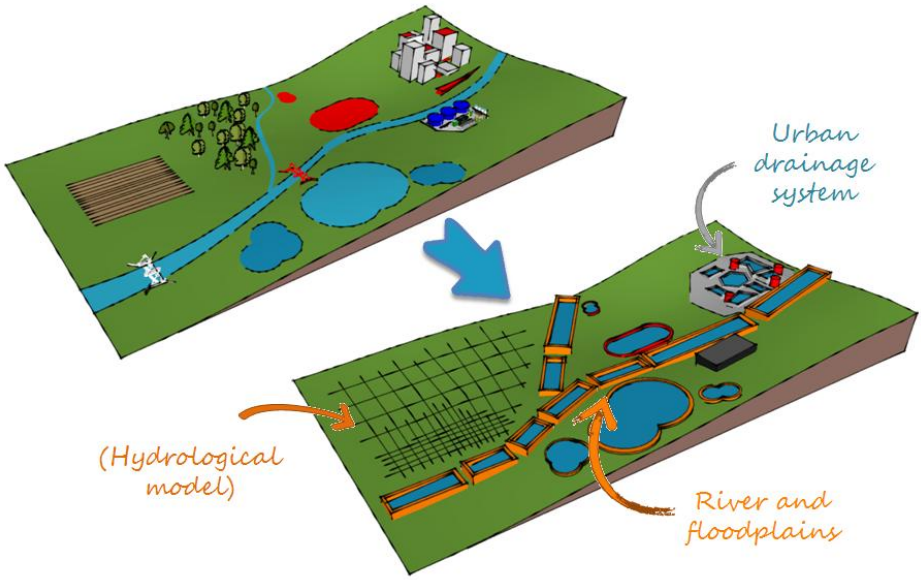


Figure 1.2: Schematization of the proposed conceptual modelling framework at catchment level.

1.4 Previous work related to conceptual water quantity modelling of rivers and urban drainage systems

In the past, doctoral researchers at the KU Leuven already investigated conceptual water quantity models.

Vaes (1999) studied the influence of rainfall and model simplifications on combined sewer system design. He stressed the importance and need for

continuous (long term) simulations, so the real variability of rainfall and antecedent conditions can be accounted for. To enable performing long term simulations in a short time frame, he developed a simple yet efficient conceptual modelling approach to calculate combined sewer overflow emissions. The approach was based on reservoirs and comprised only three main parameters for flow routing, namely the storage, throughflow and concentration time of the system. In addition, the total storage was divided in static and dynamic storages, and an extra storage component during overflow events. He already recognized the need to account for instantaneous relationships, instead of using solely maximum values, and applied a piecewise linear parameterization of these relationships. The approach was incorporated in the modelling system Remuli. Although the approach is useful for many applications, it suffers from several fundamental shortcomings. It is, for instance, not suitable of modelling variable backwater effects, or pressurized and reverse flows. This severely limits model accuracy. In addition, the flexibility in terms of model structures and topology, and integration capabilities are also restricted.

Different researchers focused on conceptual river modelling. **Timbe** (2007) developed a first conceptual modelling approach based on the classical reservoir-type topology that emulates the results of detailed full hydrodynamic river models. Flow routing was based on linear reservoir theory, while floodplains were connected to the river using simple weir-type equations. Water levels were calculated using linear equations. Different Master theses (Van Steenbergen and Verbelen, 2009; Swennen and Vreys, 2010; Vissers and Wolfs, 2011) elaborated this approach to conceptualize the Dender River. **Villazon** (2011) expanded the flow routing component to transfer functions (TFs), and added a non-linear transformation (NLT) to account for floods without the need of explicitly modelling fluxes to and from floodplains. He applied the approach to a part of the Dender River in Belgium, and the Piraí River in Bolivia.

The use of simplified river models for model-based optimization of hydraulic structures for flood control was also studied. **Barjas Blanco** (2010) and **Breckpot** (2013) approximated the de Saint-Venant equations using a linear model for every reach. They solved a quadratic programming problem to obtain the optimal flows at locations near weirs given certain constraints, set-points and rainfall predictions, and consequently calculated the corresponding gate positions. Due to the complexity of the problem, they only applied their approach to a very small part of the Demer River at the inlet of the Schulensmeer. **Chiang** (2015) used an alternative conceptualization strategy which is very similar to the

approach introduced by Timbe (2007), but incorporated weir equations explicitly in the model for flow routing. He used genetic algorithms in a model predictive control setting to obtain the best gate regulations in real time for a larger part of the Demer River.

All these studies did not primarily focus on the conceptual modelling of rivers specifically, but rather used conceptual models as a means of reaching other objectives. Consequently, additional research is required to address several critical shortcomings of the used preliminary conceptual modelling approaches. Firstly, the conceptualization procedure is not formalized and no generic approach exists, forcing modellers to rely on self-programmed algorithms and to determine how to conceptualize the system on their own. Logically, this leads to inaccurate models and possibly even to erroneous simulations. Secondly, and more fundamentally, the developed approaches are too simplistic. All listed researchers used very similar and conventional model structures for flow routing and the simulation of other states, which lack powerful emulation capabilities to capture all important and significant effects. In particular, aggregating processes on different scales, which is needed to focus on dominant mechanisms as discussed in §1.2.1, can be very challenging since more and more factors affect the investigated process. The incorporated model structures should be flexible enough to adapt themselves to the desired level of aggregation and emulate reality closely. Conventional linear reservoir theory and derivates are not appropriate for such emulation modelling. The limited number of incorporated model structures can lead to significant model deviations. Thirdly, the models have been used for only a few applications and not yet for scenario investigations. It is unclear if and how the conceptual models can be employed for such analyses. Fourthly, conceptual model can, just as the detailed models, suffer from numerical instabilities. Since even no fixed numerical calculation scheme exists for the conceptual models, model stabilization techniques have not been investigated at all. Finally, and arguably the major reason that impedes a broad uptake of conceptual modelling in general, is the lack of user-friendly and effective calibration software and the resulting complexity of configuring conceptual models. As stated earlier, the modellers have to decide themselves how to delineate the conceptual model topology, which algorithms to use, how to handle all calibration and validation data, and how to assemble all elements in a model suitable for performing simulations.

1.5 Objectives

Based on the proposed framework for developing optimal water management strategies and the needs identified in the literature review summarized in the previous sections, the following three main objectives were defined: (i) development of a new efficient and accurate conceptual water quantity modelling approach for rivers and sewers, (ii) development of an accompanying semi-automatic model structure identification and calibration tool and (iii) application of the developed modelling approach and tool on various selected case studies.

The first objective aims at **developing a conceptual water quantity modelling approach** applicable for rivers and sewer systems that is employable in the proposed decision support system for optimal, integrated and future-oriented water management. The modelling approach should be developed from a pragmatic perspective: the aim of the approach is, as ever, to give a realistic representation of the considered system, while ensuring that the structure and algorithms are tailored to specific technological needs and applications in water management. As identified in previous sections, the model should thus emulate reality accurately, while striving for parsimony and simplicity. The requisite high computational efficiency and the ability to simulate events in a very short time span are crucial in the development of the modelling approach. Other key requirements are described in §1.2.1. Existing conceptual models mainly lack accuracy, flexibility and/or suffer from a difficult and time-consuming configuration, and therefore do not meet the outlined requirements. To achieve this objective, the following tasks were identified:

- Inventory, apply, improve and expand different existing conceptual model structures and techniques for identification and calibration. Besides more conventionally used structures in hydrology such as reservoir-type structures, special attention should be paid to powerful upcoming modelling technologies, such as neural networks and other self-learning models. The selected structures must be employed in experiments at subcomponent level and be evaluated for aspects such as accuracy, stability, robustness, ease and speed of configuration, and extrapolation behavior. If needed, the structures are expanded and improved.
- The model structures must be integrated in a modelling framework. This requires adjusting and generalizing the investigated model structures, and

formalizing the conceptualization procedure and model topology. The interaction of different model structures and the error propagation during simulations are examined.

- The numerical stability during simulations and model robustness to structure or input changes are investigated. Different techniques for model stabilization are analyzed and applied.

The second objective encompasses the **development of a semi-automatic tool for conceptual model configuration**. Setting up conceptual models is a time consuming process requiring expert knowledge. Therefore, a software tool must be developed based on the created modelling approach (objective 1) that supports the configuration of conceptual models in a step-wise manner. The tool must assist the user in defining the conceptual model topology, and propose appropriate model structures and parameter sets, which can be accepted or adjusted by the modeller. This approach also allows including expert knowledge of the modeller. Finally, the configured model components must be gathered fully automatically by the software in a model script.

Most rivers and sewer systems in Flanders are already modelled in hydrodynamic software. To ensure maximum use is made of such existing models, an interfacing must be foreseen with the widely used InfoWorks and MIKE software packages to import simulation results and relevant network parameters. The software must be equipped with graphical user interfaces (GUIs) to improve user-friendliness. Future expansion of the software and interfacing with other software components, such as for conceptual water quality modelling, must be possible.

The third objective encompasses the **application of the developed modelling approach and software** on several case studies. This ensures that both are applicable for a wide range of applications and the models meet the requirements for future-oriented and optimal water management. Multiple case studies were carefully selected involving both river and sewer models with varying response characteristics, different analyses and objectives. Section 1.7 elaborates on the elected study areas.

1.6 Chapter by chapter overview

The doctoral dissertation can be divided in two main parts. The first part, comprising Chapters 2 to 5, presents the research on different model structures,

algorithms, the developed conceptual modelling approaches for rivers and sewer systems, and accompanying software tools. These chapters address the first two objectives as defined in §1.5. The literature overview and analyses of model structures and algorithms are not described in separate chapters. Instead, the developed conceptual modelling methodologies for rivers and sewer systems are presented in Chapters 2 and 3 respectively, which are the result of the extensive research on model structures and algorithms. These two chapters form a synopsis of the most fundamental developments of the doctoral research and are therefore presented first. The chapters have a very similar structure, facilitating comparison of both methodologies. Next, Chapters 4 and 5 focus on two specific modelling problems, namely the modelling of floodplains and rating curves. Additional model structures and modelling approaches are presented, that can be used as an extension to the general conceptual modelling methodologies. Each chapter in this part contains an application of the presented work on a case study to demonstrate the features and functionality of the research.

The second part of the dissertation, consisting of Chapters 6 to 8, focuses on several applications in water management. By elaborating these case studies, the use, applicability and potential of conceptual models that are configured according to the developed approaches are illustrated. These chapters address the third and final objective of the doctoral research.

Figure 1.3 provides a schematic overview of the different chapters and their relation to the three predefined objectives. The following text summarizes the content of each chapter.

Chapter 1 articulates important requisite model characteristics to comply with the needs of future water management. Next, a classification is given of different model types for water quantity modelling, together with a brief discussion of past, present and possible future use of the identified model types. A generic modelling framework is proposed towards future-oriented, sustainable and integrated water management. Based on research needs, three main objectives were defined for the doctoral research.

Chapter 2 presents the newly developed conceptual modelling approach for rivers. First, fundamental modelling principles are discussed, together with the general conceptualization procedure of rivers. Different model structures incorporated in the modular methodology are presented. In addition, a new innovative discrete solver based on the Runge-Kutta family is introduced. This

solver uses an adaptive time step to ensure model stability and minimizes computational effort during simulations. Next, the developed accompanying semi-automatic software tool, named Conceptual Model Developer (CMD), is presented. The conceptual modelling approach and tool are demonstrated on a case study, in which a detailed full hydrodynamic InfoWorks RS model of the Marke River is transformed into a conceptual model.

Chapter 3 discusses the developed conceptual modelling approach for sewer hydraulic simulations. The methodology reasons from the same modelling perspective as the approach presented in Chapter 2, but is tailored to urban drainage systems. Consequently, a modified model topology and other model structures are employed. These changes and the general content of the approach are documented extensively. Again, a software tool was developed. This tool has the same name as the tool to model rivers, but is coded independently, although both tools share basic programming scripts and functions. In contrast to the modelling approach for rivers, a simple rainfall runoff module is incorporated, since calculating the runoff and other input series is highly intertwined with the definition of the model topology. Finally, a case study is treated by conceptually modelling a sewer system near Ghent based on simulation results and data of a detailed InfoWorks CS model.

Chapter 4 presents an alternative and powerful method for lumped floodplain modelling using Adaptive Neuro Fuzzy Inference Systems (ANFIS) with Takagi-Sugeno structures in a mechanistic setting. This approach originated from the research on different model structures and algorithms. The flow dynamics of a floodplain and river-floodplain interactions can be very complex. Due to the combination of several state-of-the-art model structures, the presented approach can emulate such complex dynamics accurately. Although the modelling approach presented in Chapter 2 can emulate the vast majority of floods and dynamics accurately, more powerful approaches can be needed when lumping processes on very large scales, or exceptionally high accuracies are required. In such cases, this approach can be used as an extension to the general methodology for the conceptual modelling of rivers. The methodology is compared to the more conventional use of fixed optimized weir equations for four different sites along the Dender River. The results show that the approach using ANFIS models outperforms the use of weir equations for the investigated case study. The approach was not incorporated in the CMD software, since thorough knowledge is required to properly configure ANFIS structures, and the CMD tool includes model structures that can adequately emulate floods in most situations.

Chapter 5 focuses on the modelling of discharge-stage relationships, also known as rating curves, that are affected by hystereses. This problem of non-univocal relationships proved to be an interesting topic to investigate different model structures, and has great potential for future applications due to the widespread use of rating curves in hydrology. Several approaches are investigated that try to emulate looped rating curves using different model structures. The performance of the commonly used single rating curve is compared to a variant that incorporates a time varying parameter, which is estimated using a state dependent parameter algorithm. In addition, two machine learning techniques are examined, namely the rarely used M5' model tree and artificial neural networks. All approaches are tested on two case studies in the Dender catchment using data of detailed hydrodynamic models. Each model's sensitivity to limited data availability, discontinuities and outliers in the configuration data, and the extrapolation performance are assessed. Naturally, this research can be used to model rating curves, but most structures were subsequently incorporated in the general conceptual modelling approach for rivers presented in Chapter 2. Their use is not limited to solely simulating flows based on water levels (as in rating curves), but they were employed in various settings.

Chapter 6 is the first chapter of the second part of the dissertation, which focuses on different applications in water management. This chapter presents a first application: the impact analysis on flood probabilities of retention basins that will be built or expanded in the future and varying gate controls of several hydraulic structures along the Marke River. The conceptual model configured in Chapter 2 was coupled with a conceptual model of the Dender River, and used to perform long term simulations. The conceptual models allowed incorporating the investigated adaptations explicitly. Due to their very short calculation time, long term simulations can be performed quickly. Such long term simulation is required to account for antecedent conditions. By statistically analyzing the simulation results of the models representing the current and future states of the rivers, it was possible to quantify the reduction of flood frequency and magnitude versus empirical return periods.

Chapter 7 presents a second application using conceptual models to generate flood probability maps for the Dender River. This river is part of the operational flood forecasting system of Flanders Hydraulics Research. The conceptual model transforms the confidence intervals of forecasted water levels and exceedence probabilities of predefined alert levels at gauged locations, which are provided by the flood forecasting system, into flood extents. Conceptual models are required

due to their computational efficiency: these probability maps must be generated in real-time, enabling emergency units to take appropriate action in time. This study also illustrates how the outcome of conceptual models can be visualized using flood maps.

Chapter 8 illustrates the proposed generic framework for integrated water management introduced in §1.3 in practice. The impact of combined sewer overflows (CSOs) is quantified using coupled river-sewer conceptual models that simulate both water quantity as quality. The conceptual quantity models were individually calibrated using simulation results of detailed full hydrodynamic river and sewer models. The models were coupled to form one simulation model, and then used to perform long term simulations. The impact on the river water quality is quantified by comparing the 90^{th} and 99^{th} percentiles with and without taking into account the extra discharges that enter the river through the CSOs.

Chapter 9 summarizes the findings of the research and articulates some future research subjects.

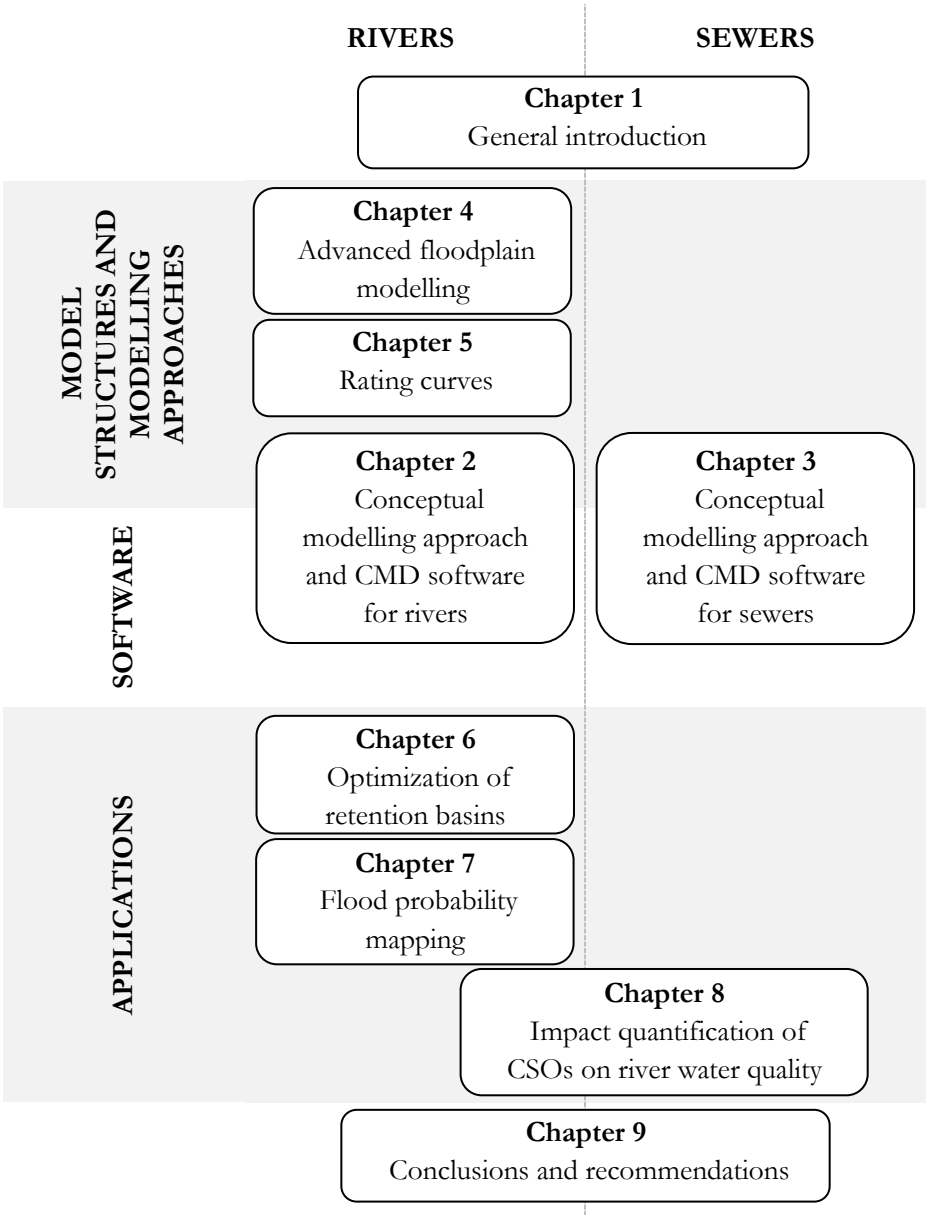


Figure 1.3: Structure of the dissertation.

1.7 Study areas

This section summarizes the different study areas and investigated applications. A more elaborate discussion of the study areas, available data and models can be found in the specified chapters or references. Figure 1.4 illustrates for which rivers and sewer systems the developed conceptual modelling approach was employed.

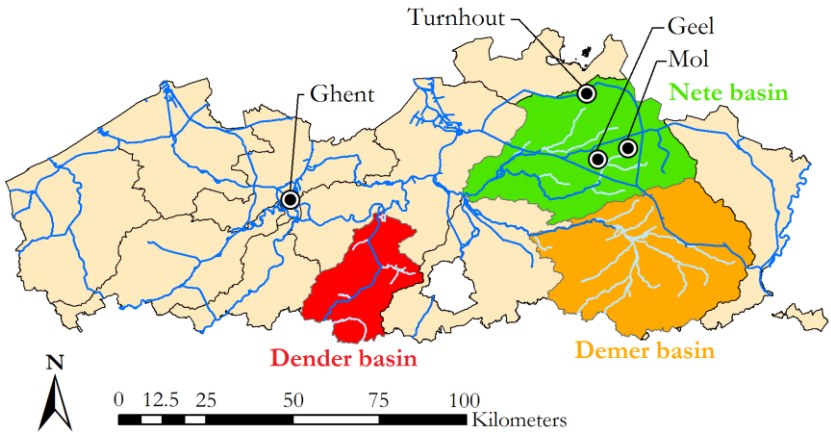


Figure 1.4: Overview of Flanders with the different VHA-zones demarcating the main river catchments, and all navigable rivers and rivers of first category (blue). Unnavigable rivers of second category of relevant basins are also marked (pale blue). The basins (rivers) and cities (sewer systems) that were conceptually modelled are indicated.

The conceptual modelling approach and software were developed by elaborating selected case studies. This allowed for identifying practical model requirements and obstacles during the early stages of model development, and ensured broad and effective applicability of the modelling approach. The modelling methodology and software were developed based on simulation results of detailed hydrodynamic models comprising:

- The Marke and Dender Rivers (Chapters 2, 4 and 5)
- The sewer system of Oostakker and Sint-Amandsberg, two districts of the city of Ghent (Chapter 3)

The developed methodology and tool were used in several applications. Three studies are described in depth in the PhD manuscript.

- Application 1: Impact quantification of flood control basins and varying operational rules of hydraulic structures along the Marke River (Chapter 6)
- Application 2: Flood probability mapping on the Dender River (Chapter 7)
- Application 3: Impact analysis of combined sewer overflows on the water quality of the Nete River (Chapter 8)

During the course of the research, the conceptual model and software were also used by other researchers for several applications. The two most extensive applications are briefly described hereafter, but are not adopted in the remainder of the text. The reader is referred to literature for additional information.

- The first study quantified the impact of source control versus end-of-pipe solutions on sewer and river floods (De Vleeschauwer et al., 2013 and 2014). The efficiency of three adaptation strategies in terms of the reduction in sewer and river flood frequencies and volumes, and sewer overflow volumes was evaluated for the city of Turnhout: (i) source control as a result of blue-green water integration (retention and infiltration in open and/or green spaces in the city center), (ii) retention basins located downstream of storm water sewers, and (iii) end-of-pipe solutions based on river flood control reservoirs. A conceptual model of the sewer system of the city of Turnhout was coupled with conceptual models of the Aa and Visbeek Rivers. All conceptual models were calibrated using simulation results of detailed hydrodynamic models. The integrated river-sewer conceptual model was used to perform a long term simulation (historic 100-year rainfall series), followed by a statistical analysis of the results. The study illustrates the importance of improving the interface between urban design and water management, and between sewer and river flood management.
- The second study used conceptual river models for real time control of hydraulic structures to minimize flood damage along rivers in the Demer basin using Model Predictive Control (MPC) and Genetic Algorithms (GA). The conceptual models play an essential role as their computational efficiency allows numerous iterations in real-time. Preliminary results are described in Vermuyten et al. (2014) and Van den Zegel et al. (2014).

2

Conceptual modelling approach and software for rivers

2.1 Introduction

This chapter describes the newly developed conceptual modelling approach and accompanying software for river systems. First, fundamental modelling principles are discussed, which are also valid for the modelling approach of sewer systems (this approach is presented in Chapter 3). Both methodologies reason from the same theory and modelling perspective, but incorporate different conceptual model topologies and structures, tailored to the type of system. The remainder of this chapter focuses solely on the conceptual modelling approach for rivers. The model topology and different incorporated structures are discussed, together with the software that incorporates the modelling methodology. A case study is described to illustrate the conceptualization procedure, the functionality of the modelling approach, and the model's accuracy and calculation speed.

The next section (Section 2.2) describes the basic modelling principles and configuration procedure. Section 2.3 describes the different model components. Since the modelling approach has a modular structure, a wide gamut of model structures is incorporated to calculate discharges and water levels. The user can select the most appropriate model structure based on the river dynamics,

available data and intended use. An optional extension is presented using neural networks to account for rapidly changing water surface profiles. Section 2.4 presents an innovative discrete solver based on the Runge-Kutta family that uses an adaptive time step during simulations to ensure stability and maximize the computational efficiency. It is possible to apply the solver solely to model components which are prone to instabilities. In Section 2.5, we discuss the accompanying semi-automatic software, named Conceptual Model Developer (CMD). This tool facilitates and speeds up the time-consuming process of model structure identification, calibration and build-up. It is closely interfaced with the InfoWorks RS software, but can handle spreadsheet data as well. It collects the required data, guides the user in a step-wise manner through the entire set-up process and proposes the best suitable model structures using advanced identification and optimization techniques for model calibration. After set-up, the different configured model components are automatically gathered and combined into a model script that represents the derived conceptual model, together with the requisite boundary data to simulate selected events. Finally, Section 2.6 illustrates the applicability and effectiveness of the conceptual modelling approach and the presented calculation scheme by investigating a case study. A conceptual model of the Marke River is built using simulation results of a detailed hydrodynamic InfoWorks RS model.

The chapter is based on the following publication:

> WOLFS, V., MEERT, P., WILLEMS, P., 2015. MODULAR CONCEPTUAL MODELLING APPROACH AND SOFTWARE FOR RIVER HYDRAULIC SIMULATIONS. ENVIRONMENTAL MODELLING AND SOFTWARE 71, PP. 60-77.

2.2 Conceptualization of the river system

2.2.1 Modelling principles

The conceptual modelling approach is developed from a pragmatic perspective: it tries to emulate reality accurately, while meeting the various model requirements outlined in §1.2.1. It is of paramount importance to assure its applicability as decision support tool for developing water management strategies and for performing various water system analyses.

Applying solely a structure-based dynamic emulation approach (Castelletti et al., 2012b), in which the mathematical structure of a full hydrodynamic model is

manipulated to produce a simpler and computationally more efficient form (e.g. the simplification from the full dynamic equation to the kinematic formulation; see also §1.2.2.1), yields models with insufficient computational improvement. Instead, a data-based mechanistic approach is applied to configure a conceptual model with different model structures using simulation results of a detailed model or measurements. To strike a balance between computational efficiency, and model credibility and accuracy, the presented modelling methodology manipulates the original model on two levels, namely the network topology and the momentum equations used to estimate the states in the system. The so-called lower-fidelity approach (e.g. Razavi et al., 2012) is used to simplify the network topology by applying the storage cell concept, in which the entire system (both the river and floodplains) is divided into distinct units. Such division can be physically based and preserve the main topology of the system, but lumps the processes in space and is less detailed than a full hydrodynamic (high-fidelity) model. Surface response surrogate relationships, which are data-driven function approximations of the true response, are derived to calculate the states of the cells and inter-cell fluxes. Flows can for instance be estimated via the more traditional approach of storage-flow relationships, but can also be related to water levels. The latter is imperative for dealing with externally controlled discharges as pointed out by Fenton (1992). Different state-of-the-art model structures are incorporated in the modelling approach, resulting in a highly flexible approach that can be tailored to the user needs and system response.

Due to the data-driven character of the conceptual modelling methodology, a large amount of data is required to ensure that all system components are identifiable. The data required to set-up a conceptual model comprises time series of flows, water levels and gate levels, but possibly also parameters and logic controls of (controllable) hydraulic structures. Numerous other studies have discussed this necessity for extended data sets (e.g. Hunter et al., 2007). Since obtaining reliable data is difficult and expensive, the use of "virtual" data in the form of simulation results of more detailed (e.g. full hydrodynamic) models is often proposed to configure conceptual models (e.g. Meirlaen et al., 2001; Vanrolleghem et al., 2005). Hydrodynamic models do not require such a large set of water level and flow measurements due to the purely physically-based equations they incorporate. The presented methodology is therefore primarily aimed at the nominal emulation of these detailed models. However, if detailed models cannot be employed for some reason, one can also use (a limited set of) measurement data to configure the model. Restricted data availability has three

major consequences on the model set-up. Naturally, the overall accuracy of the model will likely decrease. Dynamic effects, such as backwater effects, that are not present in the data set, cannot be predicted due to the data-based character of the approach. Secondly, the modeller is confined in its use of model structures to predict variables. Some structures rely on parameters without direct physical meaning and require comprehensive data sets to estimate (see §2.3 for an overview of the incorporated model structures). Thirdly, the model detail (i.e. the level of detail in the network topology) is limited by the available data. Fluxes need to be calibrated between the different elements constituting the topology, which require either data to estimate, or hypothetico-based parameter estimations. The latter is, however, not always straightforward or even possible.

2.2.2 Model configuration procedure

The configuration of the conceptual model starts with defining the desired network topology. Since the storage cell concept underlies the proposed methodology, the river and floodplains are transformed into multiple interconnected cells. All processes are lumped in space and the defined cells do not have to form a regular or even Cartesian raster, but can cover irregular polygonal areas. The extent of each cell, and hence also the degree of agglomeration, depends on the intended use, the desired accuracy and the behaviour of the river. It is therefore not possible to fully automate this process, and hence the user must define the system topology manually using template spreadsheets (see Appendix A). These template spreadsheets enable the developed modelling software to interpret the desired network, collect the required data and build the model (see also Figure 2.1 and §2.5). Based on experience, some qualitative guidelines where cells are usually demarcated can be given. In general, the river network will be divided at (controllable) hydraulic structures and other elements that can influence the flow significantly or cause (variable) backwater effects, such as bridges and constrictions. Naturally, dikes and embankments have a significant impact on the river's dynamics and are therefore usually incorporated as cell boundaries. The river is also partitioned just downstream of confluences and bifurcations to allow for flow estimation in each branch. Finally, in areas where model accuracy is of paramount importance, one can consider to define smaller cells. This reductionist and lumped approach has both supporting and opposing arguments. It is beneficial that the modeller is forced to analyse the river dynamics and to think about to which scale processes need to be simulated explicitly. By lumping processes and thereby reducing the

number of model elements, the overall complexity of the conceptual model lessens and consequently, the required calculation time decreases. Naturally, this entails intrinsically a loss of accuracy on the fine-scale level, including distributed patterns of turbulence, velocity and acceleration, but the developed methodology is not designed to account for such dynamics at all. Instead, the simplified equations and modules are constructed to give a good representation at reach scale. In addition, it is likely that uncertainties in the representation of roughness and topography will influence the results more than potential inaccuracies in the hydraulic representation at smaller scales (Molinaro et al., 1994). Thus, it is up to the modeller to find the optimal balance between simplifications and the required accuracy.

Next, the model structures have to be identified and calibrated. This is a complex and time-consuming task, due to the use of different techniques and data requirements. Therefore, a software tool (Conceptual Model Developer, CMD) was developed. This tool automates many of the time-demanding tasks. Section 2.3 describes the different available model structures incorporated in the methodology and software, while §2.5 elaborates on the functions and features of the software tool.

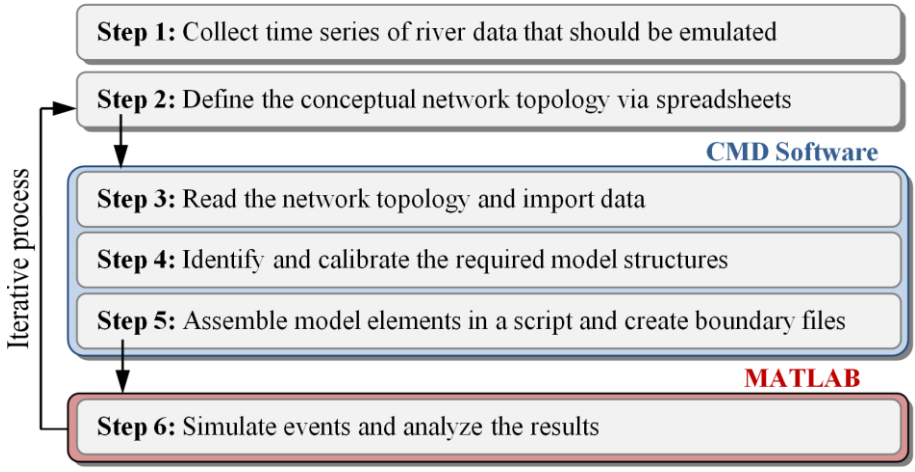


Figure 2.1: Different steps of the set-up procedure of the conceptual models.

The three main stages for model set-up identified in the beginning of this section can be split up in the 6 steps depicted in Figure 2.1. This figure synthesizes the workflow of setting up conceptual models using the presented methodology and software. The CMD software guides the user through steps 3-5, while simulations are performed in the MATLAB environment (see §2.5.1 for further details).

2.3 Model components

The presented modelling approach is based on the principle of modular and mechanistic design. Therefore, the modelling framework comprises different model structures for predicting flows, volumes, water levels and gate positions of controllable hydraulic structures. Depending on the (nonlinear) behaviour of the river, the modeller can pick the most suitable and effective model structure to emulate a state. These different subsystems can be freely connected and combined in a conceptual model, without the need of introducing additional relationships or transformation modules. The interfacing of different subsystems is done automatically using the CMD software (see §2.5.4). Table 2.1 summarizes the incorporated model structures, their inputs and application domain. The following sections elaborate on each of these model structures.

2.3.1 Flow calculation

Calculation of the flow forms the core element of a conceptual river hydraulic model. The approaches to calculate the discharge can be categorized into three distinct groups, namely based on (1) water levels (see §2.3.1.1 and 2.3.1.2), (2) storage (§2.3.1.3 and 2.3.1.4) and (3) an artificial set of rules or a Programmable Logic Controller (PLC, §2.3.1.5). The methods that rely on water levels are applicable when there is no linearity between rainfall excess or upstream inflow and the stream flow response. More specifically, this makes them capable of emulating backwater effects and tidal influences between different cells.

2.3.1.1 Option 1: Weir equations

The first and probably most obvious manner to calculate the flow based on water levels is by applying a selected set of weir equations. Such equations are deemed to describe the flow over a specific fixed or movable hydraulic structure, such as a weir, sluice, culvert, spill or orifice. Numerous formulations can be found in literature for each of these structures (e.g. Henderson, 1966; Liggett and Cunge, 1975; Swamee 1992; Yen et al., 2001). This study uses the discharge equations implemented in the full hydrodynamic InfoWorks RS software, which were either derived and tested on a physical model, or adopted from literature (see Innovyze (2014a) for further details). Most of the equations differentiate free and submerged (drowned) flow. One can simply adopt a set of parameters if known a priori, but it is also possible to identify the optimal set given some flow and stage data.

Table 2.1: Overview of the different modules, their inputs and usage or advantages. Symbols: Q = flow; V = volume; WL = water level; GL = gate level; HS = hydraulic structure; ANN = artificial neural network; TF = transfer function; NLT = non-linear transformation; PWL = piecewise linear relationship; PLC = programmable logic control; RC = rating curve; HC = hypsometric curve; M5' = M5' model tree; WSP = water surface profile.

Output	Type	Model structure	Inputs	Usage or advantages
Q	HS equations	-	WL_{up}, WL_{down}, GL	Explicit modelling of hydraulic structures
	ANN	ANN	WL_{up}, WL_{down}	Emulation of complex flow dynamics
	Storage-flow	PWL	V	Simplest flow regimes, transparent, computationally efficient
	TF (and NLT)	TF (and PWL)	Q	Natural flow regimes without variable back water effects
	Abstractions	PLC	Various states	Flow influenced by controls or manual operations (e.g. pumps), infiltration modules
V	Continuity equation	-	V, Q	Closes water balance in each cell
WL	RC	PWL, M5'	Q, $\Delta Q^{(1)}$	Default water level prediction method. (1) ΔQ is only applied for M5'.
	HC	PWL, M5'	V, $\Delta V^{(2)}$	If the use of RC is infeasible (e.g. when the flow is controlled by movable hydraulic structures). (2) ΔV is only applied for M5'.
	Rotating WSP (3)	ANN	Various states	Correct WL predictions for dynamic effects. (3) extension to HC and RC
GL	-	PLC	Various states	Estimation of gate positions of movable hydraulic structures

Two different optimization algorithms are implemented in the CMD software for the optimization of parameters of the weir equations. The first is the Shuffled Complex Evolution (SCE-UA) algorithm (Duan et al., 1992; Sorooshian and Gupta, 1995). This is an efficient and robust global optimization algorithm with widespread use in hydrological applications and was later on further improved (see SCEM-UA: Vrugt et al. 2003a; 2003b). The second tries to determine the best parameter set via a constrained interior-point optimization (see Byrd et al. (2000) and Waltz et al. (2006) for more information), which is faster in convergence than the SCE-UA algorithm, but could result in local optima. We recommend using the interior-point optimization by default due to the much faster convergence towards a solution. If the results are unsatisfactorily and the optimization algorithm may have found a local optimum, the modeller can use SCEM-UA. Optimization termination is determined based on several tolerance criteria involving the coefficient of determination R^2 between the fitted and original time series of flow.

2.3.1.2 Option 2: Artificial neural networks

The rate of flow in rivers and floodplains is influenced by many natural phenomena. The flow over embankments in particular is a very complex matter due to the often irregular topography and different vegetation covers (and hence resistance forces). It is possible that weir equations fail to adequately emulate the flow process due to their rigid model structure and confined flexibility. Naturally, even the most advanced optimization algorithms cannot overcome such limitations and consequently alternative model structures should be explored. Soft computing models are due to their flexible nature and learning capabilities well qualified for these problems. Artificial neural networks (ANNs) in particular can deliver more accurate results compared to optimized weir equations when emulating the flow between a river in and adjacent floodplains, as demonstrated by Wolfs and Willems (2014a) for a Belgian case study. The ANNs are driven by a selected up- and downstream water level, similar to the weir equations. Different ANNs can be constructed using the early stopping technique or a Bayesian regularization framework (as proposed by MacKay, 1992) to ensure adequate generalization capabilities of the obtained networks. Training of the ANN is stopped in the early stopping technique when the error on validation data increases during a selected number of subsequent epochs. Bayesian regularization minimizes, in addition to the squared errors as done in early stopping, also the weights in the ANNs. Next, different ANNs can be combined

via averaging to yield an ensemble ANN with enhanced generalization. Due to their adaptable structure and data-driven nature, the obtained ANNs can outperform the optimized weir equations, especially when aggregating processes on a larger scale.

2.3.1.3 Option 3: Direct storage-flow relationship

In some situations when no backwater effects prevail and the flow is not externally controlled, it is possible to identify a direct relationship between flow and storage. In such straightforward cases, the relationship between both can be captured by simple piecewise linear equations. Hence, this approach is virtually equal to the widely adopted linear reservoir concept and avoids the optimization or identification of more complex structures.

2.3.1.4 Option 4: Transfer function (TF) and optional non-linear transformation

Simple direct relationships between storage and flow will often not deliver the desired results. Elaborations of the concept of a relationship between storage and throughflow have been developed (e.g. Nash, 1957; Dooge, 1959; Diskin, 1964; Kulandaiswamy, 1964 and see §1.2.2.3). It is common practice to combine multiple linear reservoirs in cascade, parallel or both to enhance the predictive power of the model. It has been proven (e.g. Jakeman et al., 1990; Willems, 2000; Beven, 2001) that such configuration can be represented by a transfer function (TF) of the following form:

$$Q_t = \frac{b_0 + b_1 z^{-1} + \cdots + b_m z^{-m}}{1 + a_1 z^{-1} + \cdots + a_n z^{-n}} z^{-\delta} \cdot I_t + \xi_t \qquad (2.1)$$

with a_1, ..., a_n and b_0, ..., b_m time-invariant coefficients, δ a general delay element, I and Q the incoming and outgoing flows respectively and ξ an error model. The backward shift operator z is defined as $z^{-1}Q_t = Q_{t-1}$. [1]

[1] This equation uses the discrete formulation of TFs. Due to their flexibility, good emulation performance and interpretability, TFs are also incorporated in the newly developed conceptual modelling approach for sewer systems presented in Chapter 3 (see §3.3.3.3). In the text on conceptual sewer system modelling, the TF theory is approached from the continuous time domain. Section 3.3.3.3 also shows the link between TFs, linear reservoir theory and the Instantaneous Unit Hydrograph (IUH) method. The use of the discrete and continuous formulation of TFs for river and urban drainage modelling respectively by the research group has grown historically.

Most of the above cited methodologies are special cases of the TF (see e.g. Spolia and Chander, 1974). These TFs are ideally suited for rivers without controllable hydraulic structures, or when there is no time-varying backwater influence. It is computationally less expensive and more robust than the presented methods based on water levels. The flow used as input in the TF equation is determined substantially by the flow coming from upstream located reservoirs, but is of course also affected by tributaries and floodplains. If the magnitudes of such side-flows are known, they can simply be summed up. However, this information often lacks in practice. In such cases, one could model each unknown flow individually using the earlier described methods and still take them into account when determining the input. Alternatively, a non-linear transformation (NLT) can be performed on the input. This combination of NLT and TF converts the flow prediction model in a "Hammerstein" non-linear model type. Such NLT is of particular interest for flows from or to floodplains, since they are most often characterized by some reciprocity. This NLT obviates the need of modelling such flows explicitly. The NLT involves a dynamic correction factor applied to the input that depends on the state of the river (usually the upstream incoming flow or storage in the reservoir) and can be identified via a state dependent parameter algorithm (see e.g. Young et al., 2001). Hence, the values of the parameters of the NLT structure vary with the value of the selected state, making it possible to capture non-linear interactions. For instance, large incoming flows can be decreased or attenuated via this NLT to imitate the flow passing to floodplains. In contrary, during the recession limb of the hydrograph, the input flow of the TF can be increased using the NLT to emulate the emptying of adjacent floodplains.

Transfer functions can be configured manually by calibrating each reservoir separately, assessing the recession constant and delay data-based (Willems, 2009), and finally transforming the obtained structures into one transfer function. Alternatively, direct model structure identification and calibration can be performed using various algorithms. Direct identification and calibration of TFs avoids prior judgment on the configuration of the linear reservoirs since no predefined order is required. After TF configuration, partial fractal decomposition of the obtained TF can provide useful insight into the system dynamics if the roots of the TF are real (see e.g. Young, 2006 and references therein for several case studies and physical interpretation of the derived systems; and §3.3.3.3). This study favours the Refined Instrumental Variable algorithm (RIV, Young, 1984) for TF set-up due to the computational efficiency,

performance and capability of assessing parameter sensitivity explicitly. The RIV algorithm is incorporated in the CAPTAIN toolbox for MATLAB® (Taylor et al., 2007; Young et al., 2007). The ideal model structure is chosen based on the R_T^2 value and YIC criterion (Young, 1989), which weighs goodness of fit against model complexity.

2.3.1.5 Option 5: Abstractions

Lastly, flows can be estimated based on a set of logic rules that depend on other states, such as the river water levels, discharges (at other locations or previous time steps), or manually controlled switches. This set-up is particularly useful for incorporating complex structures, such as pumps or infiltration modules. The rules themselves can also contain other flow calculation mechanisms (e.g. §2.3.1.1 to §2.3.1.4).

2.3.2 Water balance

Volume continuity is ensured by closing the water balance explicitly. This is done by applying the discrete form of the continuity equation and taking all incoming and outgoing flows of a storage cell into account. Note that it is unnecessary to calculate the continuity equation if a transfer function is applied to calculate the flow that already uses the sum of all discharges of a storage cell as input and has a steady state gain of unity.

2.3.3 Water level calculation

2.3.3.1 Model structures

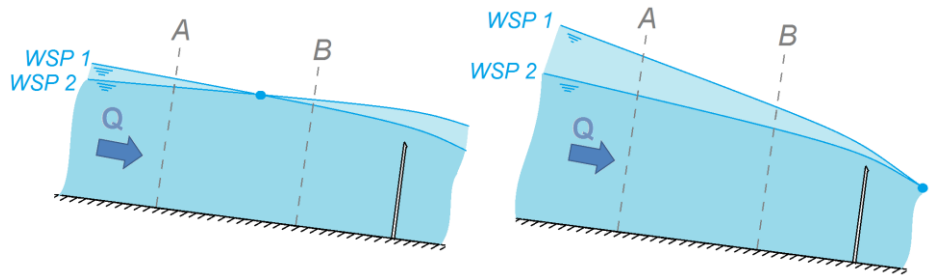
It is possible to relate the water level to different states of the river, namely the flow, storage or a water level at a different location. Depending on the chosen state, respectively rating curves (RC), hypsometric curves (HC) or direct stage-stage relationships (HH) are defined. Of course, the configured relationship of each water level is location specific.

Although the state dependencies differ, the same model structures are available to parameterize these relationships, including straightforward piecewise linear (PWL) equations and the more advanced M5' model trees (see Figure 2.3). PWLs are equivalent to one-dimensional lookup tables. Hence, they are configured by defining a set of coordinate couples. Linear interpolation is used to calculate intermediate values. The incorporated M5' model trees use in addition to the

selected state also the change of the state between two consecutive time steps. This enhances flexibility of the model element, but also increases its sensitivity to instabilities. M5' model trees share their structure with the more conventional decision and regression trees, but the leaves consist of linear functions instead of discrete class labels. Quinlan (1992) laid the foundation for the M5 trees, which were later on improved by Wang and Witten (1997) and since then denoted as M5'. The reader is referred to the latter publication for further details on the optimization procedure. Jekabsons (2010) incorporated this procedure in a MATLAB toolbox. No prior assumption of the underlying relationships between the explanatory variables and response is required. In addition, their simple if-then structure in combination with straightforward linear relationships makes them easy to interpret. M5' have been widely applied to other water resources and environmental modelling problems (e.g. Bhattacharya and Solomatine (2005), Elshorbagy et al. (2010) and Galelli and Castelletti (2013)).This approach can also emulate hysteretic behaviour as illustrated for several case studies in Chapter 5. Section 5.4.4 provides a more in-depth discussion of the use and configuration of M5' model trees.

2.3.3.2 Extension for non-univocal water surface profiles

The model structures discussed in the previous section determine the water level based on one state variable (and additionally the difference of this state variable over two consecutive time steps as second input in case of the M5' model structure). More specifically, such structure depends on the discharge (in case of an RC), the volume (HC) or a water level at another location (HH). If a univocal relationships exists for the investigated calibration and validation set, such simple set-up succeeds in emulating the water level perfectly. It is important to note that these model structures do not impose a constant water depth over a modelled river reach (represented by a single cell in the conceptual model). Similarly, the modelled water level does not have to be parallel to the bottom slope of the reach. In fact, the model structures can account for curved water surface profiles (WSP) as long as there exists an (approximate) univocal relationship between the sought water level and a state variable of the system. Figure 2.2b illustrates this principle for a theoretical and strongly simplified example. The water levels are calculated at sections *A* and *B* using separate model structures. The WSP rotates around a point outside of the investigated river reach. It is likely that for both *WSP1* and *WSP2* a univocal relationship exists between, for instance, the water levels at locations *A* and *B* and the discharge or volume of this reservoir.



(a) (b)

Figure 2.2: Schematization of various rotating water surface profiles (WSP). The end of the river reach is indicated by the schematized weir. Separate model structures are created to calculate the water levels at locations *A* and *B*. Rotation point inside the river reach itself which will likely lead to non-univocal behaviour and thus necessitating a correction module (a) and downstream of the investigated river reach, possibly leading to a simple univocal relationship (b).

In practice, the water level depends on many more factors than solely one system state. The water level and surface profile is also determined by, amongst others, the slope of the channel and the flow at a specific location. Consequently, such univocal relationships that can be emulated by the simple structures described in §2.3.3.1 do not exist in all situations. Therefore, an extension to the aforementioned model structures is required to account for such dynamic behaviour. A theoretical example that requires such correction module is shown in Figure 2.2a. One can easily see that the volume in the river reach for *WSP1* and *WSP2* are approximately equal, while the water levels in *A* and *B* differ. Hence, the relationship between the volume and water levels in *A* and *B* will show hysteretic behaviour. However, if this reservoir is controlled downstream by a moveable hydraulic structure which gate opening or level depends on the water level at locations *A* or *B*, the modeller can only resort to HC-type curves (i.e. a water level – volume dependency). Thus, the ordinary PWL model structures (see §2.3.3.1) cannot capture such relationships. In case of more complex hysteretic effects, it is possible that even the M5' model trees cannot emulate those dynamics accurately, and consequently, an additional correction module is necessary. The modelling of hysteretic effects using M5' model trees and other time varying model structures is discussed in-depth in Chapter 5.

To handle the problem of rotating water surface profiles that lead to non-univocal relationships between the water level and a single elected variable, an additional module is introduced in this section. Such hysteretic behaviour can

have numerous causes, such as tidal influences, controllable hydraulic structures or dynamics effects. Therefore, this module must be very flexible. Not only the additional model structure itself needs to be adaptable, but also the used input set. Note that reaches without controllable structures or strong variable backwater effects usually do not show such complicated behaviour, and consequently, the basic model structures already succeed in predicting the water levels accurately.

In the first phase during the configuration of the water level calculation component, one of the aforementioned model structures (PWL or M5' model tree) for one of the possible relationships (RC, HC or HH; see also §2.3.3.1) is identified and calibrated. The PWL structures ignore the hysteretic effects completely, but can already give an indication of the (approximate steady-state) water level. Next, a time varying correction factor is estimated to tackle the hysteresis. This factor is calculated by comparing the outcome of the earlier calibrated model structure (PWL or M5' model tree) with the calibration data. The correction factor can depend on one or more states, such as (part of) the incoming flow or outgoing flow, the storage of one or several cells, or a combination of these. Combinations can be made of all state variables of the system. Indeed, the input set must be this flexible, since it is impossible to enumerate a priori all possible situations and driving forces that cause the hysteretic behaviour. Hence, a data-based set-up is suggested, which involves the use of time series that exhibit such dynamics for the module's identification and calibration. As discussed earlier, ANNs are well suited to identify and parameterize such dependencies and are therefore used as model structure for this correction module. This approach also implies that it is the task of the modeller to select the most suitable input set via a trial-and-error procedure. In many cases, this will be the incoming flow of the cell. The developed software tool (see §2.5.3) facilitates and speeds up the investigation of different input sets. Secondly, due to the data-driven character of the additional correction module, only dynamic effects that are already incorporated in the calibration set can be emulated. It is very likely that fundamentally different behaviour, such as the widening of a hydraulic structure that regulates the outgoing flow, cannot be emulated correctly.

Figure 2.3 summarizes the functioning of the set-up during simulations. For every location in the river where water levels are being calculated, such scheme is configured. First, the water level is calculated via a PWL or M5' model tree. One and two states are being used respectively for this PWL and M5' model tree, as

discussed in §2.3.3.1. Next, the optional correction factor is being calculated as the outcome of a trained ANN that depends in turn on one or two states in the system. Note that this factor is only being calculated if the module has been configured for this specific location. Finally, the simulated water level from the PWL or M5' model tree is multiplied with the obtained correction factor to yield the final water level prediction at the desired location.

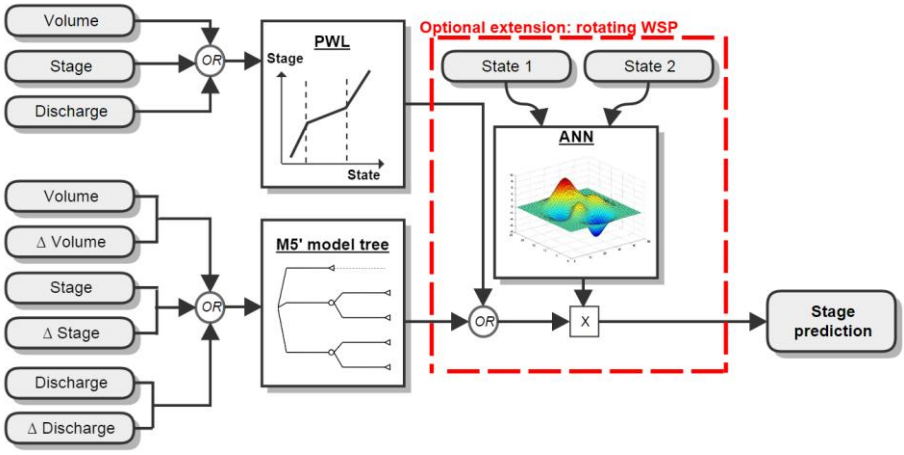


Figure 2.3: Presentation of the different modules that estimate the water level at a specified location with an optional component that accounts for water surface profiles with time varying slope.

This methodology has the advantage that the water balance is still closed explicitly, since only the water level calculation is being affected. Secondly, this module is a rather transparent yet powerful extension to the aforementioned model structures. It is possible to emulate many complex and dynamic effects due to the self-learning and flexible character of the ANNs. The outcome (i.e. the time varying correction factor) of such an ANN can be estimated a priori for several input sets and can be confined during simulations. Hence, it is possible to assess the response of the extension under varying conditions, including extrapolation. Furthermore, the extrapolation of the core model structures, either a PWL or an M5' model tree, can easily be assessed and manipulated during configuration.

2.3.4 Gate level calculation of controllable hydraulic structures

The final category of calculation modules is designed to determine the gate opening or level of controllable hydraulic structures along a river or floodplain. In practice, gate openings are usually calculated by a fixed set of logical rules, similar to a so-called abstraction that determines the flow (see §2.3.1.5). Hence, the gate opening can depend on flows, water levels, or gate openings of other hydraulic structures. In contrast to the flow abstractions, the gate opening is confined to a minimum and maximum position, and can only move at a limited speed. Such constraints can also be taken into account.

2.4 Calculation scheme and code

The final phase of model construction comprises the set-up of the calculation scheme and programming code. Due to the modularity of the modelling approach, it is in most cases very difficult or even impossible to find an analytical solution of the set of equations representing the conceptual model. Hence, a numerical solution method should be implemented.

The temporal variations in a river system have different time scales for different components. The system can react or change very quickly over time, but solutions vary only gradually during most of the simulation horizon. Such systems are often referred to as *stiff* systems in numerical mathematics. To find solutions efficiently in such settings, implicit numerical schemes are usually employed. The commercial MIKE11 software, for instance, uses an implicit 6-point Abbott-Ionescu scheme to calculate the dynamics in river systems. Such implicit schemes obtain solutions by solving a set of equations that involve both the current state of the system and the sought state. However, such implicit schemes suffer from several disadvantages. Firstly, the implementation of an implicit scheme is considerably more difficult than explicit methods (Amitai et al., 1998). Especially given the modularity of the modelling approach and the required component interoperability (e.g. in terms of adding new elements for scenario investigations or integrated analyses), the implementation of explicit schemes is significantly easier (e.g. Barros, 2007). Moreover, explicit schemes are simpler to adjust afterwards and easier interpretable by the end-user of the model. Finally, the use of implicit fixed-step solvers can be problematic in the context of real-time simulations, since the number of requisite iterations to find a solution is theoretically unbounded (Thiele et al., 2014), although different techniques are

being developed to overcome this issue by minimizing and limiting the number of iterations (e.g. Elmqvist et al., 2002). Due to these disadvantages of implicit schemes, the use of explicit solvers is investigated.

The main downside of such explicit solvers is that they are only conditionally stable. Therefore, chequerboard oscillations (e.g. Cunge et al., 1980; Hunter et al., 2007) can quickly originate and affect the entire model if the selected time step is too large during simulations. Different solutions are proposed in literature to obviate instabilities. For instance, flow limiters can be employed that confine the maximum discharges. Such approach is rather simplistic and can significantly alter the behaviour of the river and floodplains (Horritt and Bates, 2001; Bates et al., 2010). Hunter et al. (2005) state that most storage cell codes use such or equivalent conditions to ensure stability. Another remedy lies in reducing the time step as is done in many continuous variable time step solvers when instabilities are detected. The optimal time step is however both space and time dependent, and cannot be selected a priori. Indeed, some model elements are more prone to instabilities, and require therefore a smaller time step than others. It would be highly inefficient in most cases to create a calculation scheme prior to a simulation with one fixed time step for the entire model that avoids all instabilities. More adjustable and thus complicated approaches are required. Hunter et al. (2005) use a time step based on theoretical considerations of model stability, similar to the Courant-Friedrichs-Levy condition for advective flows (see also Beffa and Connell, 2001; Bradbrook et al., 2004). Such constraints are difficult or even impossible to formulate analytically for the modelling approach presented in this chapter due to its modular design. Theoretical time step limitations can only be deducted with acceptable effort for calculation nodes on a Cartesian raster and when simple momentum equations are employed. Hunter et al. (2005) apply a time step which alters during the course of a simulation, but is still fixed in space at each time step. Secondly, the optimum stable time step reduces quadratically with decreasing grid size, which leads to inacceptable simulation times for storage cells with small resolutions (e.g. Bates et al., 2010).

Consequently, the need for an efficient solution to prevent or at least confine instabilities using straightforward discrete calculation schemes becomes apparent. A simple three-phase solver is presented that does not require the analytical formulation of a maximal allowable time step or cell size to tackle the instability issues. The solver is particularly efficient if the model comprises one or more cells that require a smaller time step than the majority of cells, since the applied time step can be both time and space specific. The user can choose to apply the

variable time step solver only on instability prone cells, and all other cells are only calculated at a fixed time step. When activated, the discrete and explicit solver performs the following three actions: (1) calculate all states of all cells with a fixed prescribed time step, (2) analyze the dynamics of the cell to detect unstable behaviour and, if the results show erratic behaviour, (3) reject the solution and reiterate with smaller time steps.

The developed three-phase scheme with variable time step falls within the emerging category of so-called asynchronous or multi-rate solvers. In these schemes, each part of the system can be calculated using a different time step, enabling fast numerical solutions (Barros, 2007). Hence, the computational domain is divided into smaller subdomains in such a way that still accurate global solutions are obtained. Such domain decomposition strategy is often used to implement (even implicit) schemes on parallel computers (Amitai et al., 1998). Different schemes are being developed under the umbrellas of asynchronous (e.g. Barros, 2007; Small et al., 2013) and multi-rate solvers (e.g. Günther et al., 2001; Constantinescu and Sandu, 2007; Savcenco et al., 2007; Schlegel et al., 2009; and references therein). These solvers have a space-specific time step that can result in fast and efficient simulations, since instability-prone components of the system can be solved separately with a smaller time step.

The newly developed scheme is similar to already existing asynchronous or variable step size solvers, but differs in three respects. Firstly, the scheme is tailored to the presented conceptual modelling approach and models. The typical demarcation of the system in interconnected cells is inherently embedded in the scheme. Consequently, the domain decomposition into several clusters is done automatically on the level of individual cells. Within each cell, many equations are interlinked, while the interfacing between different cells is limited. Figure 2.4 shows a typical conceptual model of a small river reach, together with the different types of variables that are being calculated. Although this topology is limited to a single reach, one can easily expand the reasoning to more complex systems that include floodplains or bifurcations. The numerical scheme calculates all variables in all cells at a specific time t with a prescribed fixed time step Δt. Next, the dynamics of specified cells are analyzed to detect instabilities. If such instability is noticed, part of the model is being recalculated with a smaller time step. Section 2.4.2 discusses more in-depth which variables are being recalculated in case instabilities are detected.

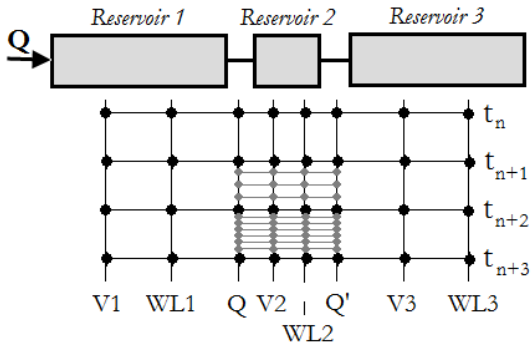


Figure 2.4: Asynchronous time stepping using the newly developed three-phase numerical scheme.

Secondly, a new instability detection criterion is formulated. Most existing variable step solvers monitor the local error to estimate the optimal step size, such as the widely used Runge-Kutta Dormand Prince solver (Dormand & Prince, 1980). This solver uses for instance function evaluations to calculate the fourth- and fifth-order accurate solutions. The difference of both solutions is then taken to be the error of the (fourth-order) solution. Other solvers with variable time steps use similar schemes. However, such approach induces significant overhead costs in the calculation. Moreover, the magnitude of the time step is purely chosen by inspecting the local error, which could still result in instable dynamics. Therefore, an alternative criterion is formulated to detect instabilities and determine a suitable time step. This criterion and alternatives are discussed in §2.4.1. The third main difference with existing asynchronous schemes is the precise formulation of the numerical solver itself. Due to, amongst others, the easier implementation and greater model scheme adaptability, a discrete and explicit solver is preferred. Therefore, an equally-weighted discrete and explicit Runge-Kutta solver is implemented that can easily deal with different time steps. The asynchronous domain decomposition is also incorporated in the formulation of the solver. This technique is elaborated in detail in §2.4.2.

2.4.1 Instability detection criterion

Instabilities result in oscillating simulated states. Hence, the problem of instability detection is reducible to recognizing such oscillatory patterns. An effective detection should not only identify such oscillations, but also fulfill the following requirements (1) performable online (i.e. "on the fly"), (2) computationally very efficient and (3) transparent and applicable with very few changeable parameters. Numerous methods exist in literature to detect oscillations in a univariate time

series, including approaches based on integrated error (e.g. Hägglund, 1995; 2005; Thornhill & Hägglund, 1997), auto-correlation functions (e.g. Miao & Seborg, 1999), spectral peak analysis (e.g. Karra and Karim, 2009), auto-regressive and moving average models (e.g. Salsbury & Singhal, 2005) and discrete cosine transformation (e.g. Li et al., 2010). Almost all mentioned detection algorithms were designed for application in the process-control industry, which involves complex processes and oscillation patterns. In contrast, oscillations in storage cell models are generally characterized by rapidly growing chequerboard oscillations with a simple recurrent pattern. These oscillations are much easier identifiable, and hence do not require such sophisticated detection methods. All of the cited methods cause substantial overhead in their computation and would involve a complex implementation, and are therefore not ideally suited for the rather straightforward instability detection for storage cell models.

Many solvers use zero-crossing detection to recognize instabilities and determine the optimal time step size. A simple alternative is proposed in this chapter based on the size of fluctuations of the simulated states around a linear trend line, instead of simple zero-crossings:

$$\theta = \left| \frac{\sum_{i=1}^{r-1} |y_{i+1} - y_i|}{r-1} \Big/ \frac{(avg2 - avg1)}{r/2} \right| \tag{2.2}$$

$$avg1 = \frac{1}{r/2} \sum_{i=1}^{r/2} y_i \tag{2.3}$$

$$avg2 = \frac{1}{r/2} \sum_{i=\frac{r}{2}+1}^{r} y_i \tag{2.4}$$

with y_i the simulated storage of a cell at time step i, r the time span which is investigated (in number of time steps) and θ the non-linearity factor indicating the magnitude of the oscillations. The trend is determined using linear interpolation through the averaged first and second halves of the investigated time span. Therefore, r must be an even number. The denominator of Eq. (2.2) represents the average change per time step along this linear trend, while the numerator calculates the average change of the simulated state per time step. The more the signal deviates from the linear path over the selected time span, the further θ deviates from unity. This criterion consists of few simple arithmetic operations, and can therefore be assessed very rapidly during simulations. Figure 2.5 illustrates the proposed criterion. The black dashed line represents the linear

trend of the simulated values, while for different oscillating time series are shown, together with their θ-factor values.

Once the θ-factor exceeds a predefined threshold value, reiterations of the examined cell will take place. The criterion assumes that the oscillating signal follows a linear trend. This is only valid for slowly evolving states. Therefore, the simulated storage is used to assess the θ-factor. In cases where the trend of the monitored state changes more rapidly and highly non-linear compared to the investigated time span r, the proposed criterion can still be used, but the most effective time span r and threshold value should be determined using trial-and-error.

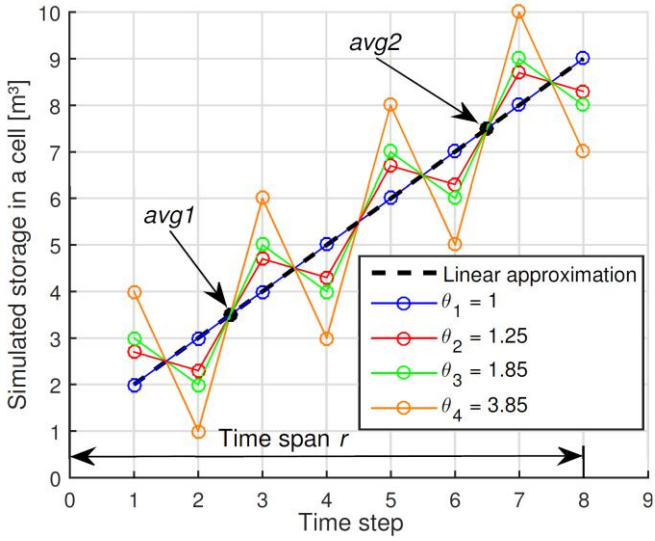


Figure 2.5: Calculation of the magnitude of oscillations (i.e. a measure of the non-linearity) of different simulated time series using the proposed instability detection criterion.

2.4.2 Variable time step solver

The third phase of the proposed solver is a discretized version of the equally weighted s-stage Runge-Kutta scheme and can be summarized by the following set of equations:

$$V_n^C = V_{n-1}^C + \sum_{i=1}^{s} k_i \tag{2.5}$$

$$k_i = \frac{\Delta t}{s} \cdot \sum_z Q_i^z = \frac{\Delta t}{s} \cdot f\left(\boldsymbol{X}^{\sim C}(t_{n-1}), \boldsymbol{X}^C\left(t_{n-1} + \frac{i}{s}\Delta t\right)\right) = \frac{\Delta t}{s} \cdot f\left(\boldsymbol{X}^{\sim C}(t_{n-1}), \boldsymbol{X}^C(t_i)\right) \quad (2.6)$$

With V_n^C the unknown storage of cell *C* at time step t_n, which is the sum of the current value V_{n-1}^C and the sum of *s* sample increments k_i. The amount of reduction of the time step can depend on the non-linearity detected in the previous phase: the larger the non-linearity θ, the more samples *s* can be introduced and hence the smaller the reduced time step will be. Each increment k_i is in turn the product of the reduced time step (i.e. $\Delta t/s$) and the sum of the estimated fluxes of the cell specified by the function *f*. Each flow Q^z is related to a vector of states $\boldsymbol{X}$ in the system, depending on the elected model structures (see §2.3.1). For instance, if flow Q^z is determined by a controllable hydraulic structure (see §2.3.1.1), it would depend on the up- and downstream water levels and gate position. All states of cell *C* itself (such as water levels and the storage), together with the z inter-cell fluxes are denoted as "internal states" ($\boldsymbol{X}^C$) and are recalculated using the reduced time step. Other relevant "external" states ($\boldsymbol{X}^{\sim C}$), such as water levels that are located in other cells but are necessary to calculate these inter-cell fluxes, are copied from the most recent prediction at t_n or t_{n-1}.

The functioning of the variable step solver is also illustrated in Figure 2.4. First, all states are being calculated at time t_{n+1} using the predefined time step size Δt based on the available states at t_n. No instabilities are detected, and consequently no iterations are necessary. Next, the states of all reservoirs are calculated at time t_{b+2}. The calculated θ-value of *Reservoir 2* surpasses a specified threshold, indicating an instability originated in this cell. Therefore, the volume (*V2*), water levels in this cell (*WL2*) and flows Q and Q' are being recalculated with a smaller time step. Note that the water levels *WL1* and *WL3* are not being recalculated, since they are located in other cells, and therefore treated as "external" states.

Hence, the proposed solver is a modified version of the original Runge-Kutta family, in which all relevant states are recalculated at the reduced time step to estimate the inter-cell fluxes. Applying a Runge-Kutta solver without this reduction of calculation nodes would be tantamount to recalculating the entire model with the smaller time step, thereby spoiling the advantage of applying a spatial variable time step. Such asynchronous and adjustable time step is an essential part of the proposed solution for obviating instabilities in a computational efficient manner. Indeed, irregular storage cell models are often

characterized by a few instability-prone cells, while most vary gradually between two consecutive time steps. The latter imposes an important condition when applying this modified scheme and puts a limit on the allowable general time step Δt. If the overall time step is chosen too large, the external states used during the reiterations will differ from their true values and this can affect the behaviour of the model. If multiple instability-prone cells are directly interconnected, changing the calculation scheme so that all these cells are recalculated simultaneously can result in improved stabilization and performance. After all cells are treated, the storages are recalculated using all obtained fluxes to ensure that the water balance remains closed.

Applying an adjustable time step also has implications on the calibrated M5' model trees and transfer functions. M5' model trees rely on the state difference between two consecutive time steps. To enable simulations with different time steps, the relevant input is scaled during simulations by multiplying it with the ratio of the time step of the training data over the simulation time step. This ensures that the conditions that determine which linear relationship is addressed and the content of this relationship itself remain unchanged. Estimated continuous transfer functions are converted to their discrete form to incorporate them in the calculation scheme. Since the parameters of discrete transfer functions are also dependent on the sampling interval, the discretization is calculated for different sampling times. All other incorporated model structures rely solely on instantaneous state predictions, and thus do not require alterations.

2.5 Software

A software tool was created to facilitate the main part of the model set-up. The tool is called Conceptual Model Developer (CMD). Before the CMD can be applied, the modeller must manually define the conceptual river topology via template spreadsheets (see §2.2 – step 2). Next, the CMD can be used to speed up the final three stages of the set-up procedure (see §2.2 steps 3-5), which are the most time-consuming tasks. Figure 2.6 summarizes the three main functions of the CMD.

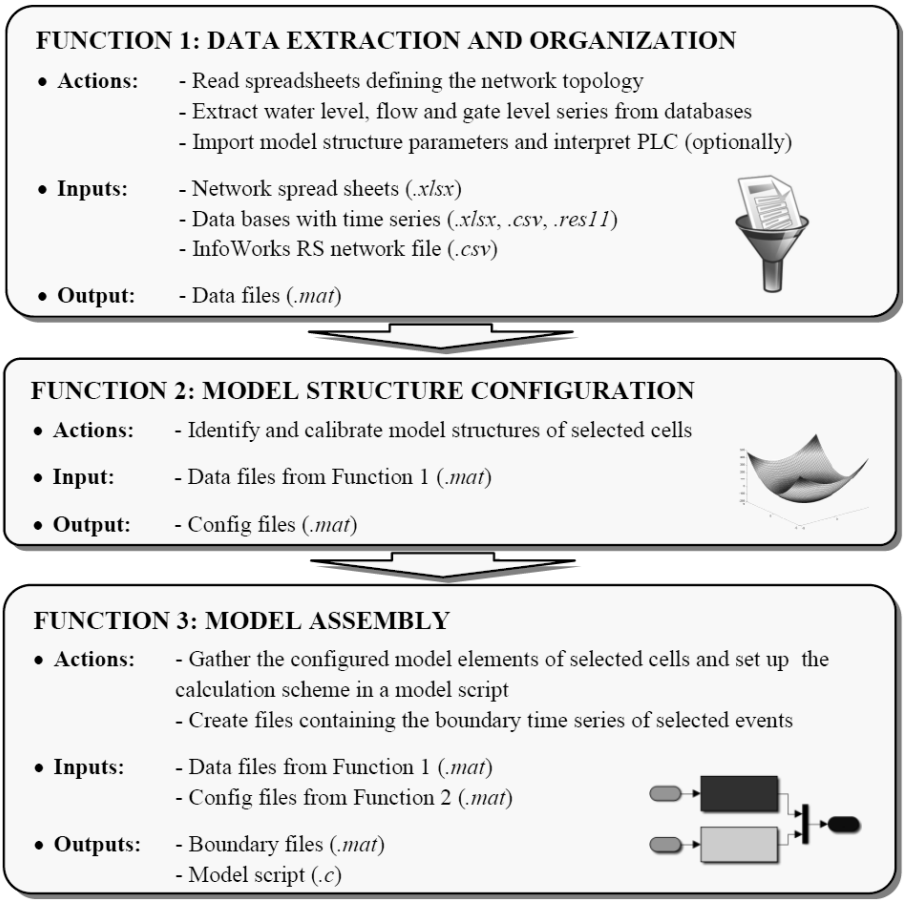


Figure 2.6: The three main functionalities of the CMD software, together with the required input files and the outcome of each function. The file formats are indicated between brackets.

2.5.1 Software implementation

The CMD software is implemented in MATLAB and equipped with a GUI. The software can be compiled in a standalone application, which means MATLAB is no longer required to run the software. In addition, all data structures that are created by the CMD tool are internally passed from one function to another. This implies that no MATLAB or programming knowledge is required to operate the CMD tool.

The calculation scheme of the conceptual model is written in a single file in the C programming language. Scripts compiled in C are computationally much more

efficient than ordinary MATLAB code. The boundary file, which contains all relevant time series to simulate events, is stored in a MATLAB structure variable. The conceptual model can be simulated in MATLAB by simply calling the model script and passing the boundary file as input argument. This means that the code can easily be implemented in other user-defined MATLAB or SIMULINK modules. Several scripts and functions were developed to visualize and interpret the simulated results, and to compare them with the original data sets used for model configuration.

2.5.2 Function 1: Data extraction and organization

When the first function is triggered by the user, the CMD tool imports the desired conceptual model topology. It is therefore required that the modeller defines this topology using predefined spreadsheets (see Appendix A). The use of such spreadsheets has the advantage that the tool can be employed to emulate any detailed network, and it is not compatible with solely one software type. Secondly, the topology is clearly structured in a transparent tabular form. Importing the network topology enables the software to learn which data is required for the model set-up and how different calculation nodes are inter-linked spatially. The latter is crucial information to ensure a fast and automated model build-up later on (see Function 3, §2.5.4). Next, the CMD tool automatically extracts all essential data from predefined databases based on the delineated conceptual network topology for selected events. This includes time series of water levels, flows and gate positions. A close integration was provided with the InfoWorks RS (IWRS) software which enables the tool to interpret and convert the content of logical rules that define gate openings or flow abstractions (see §2.3.1.5 and §2.3.4) into new formulations that can be used directly in the final conceptual calculation scheme. In addition, it can also extract other essential parameters from the IWRS network file, such as those of hydraulic structures that are modelled explicitly and gate opening limitations, including the maximum movement rate and most extreme allowable positions. Although the CMD is tailored to IWRS models, it can also be used in situations where such model is unavailable. Finally, the information is stored in several MATLAB data structures that will be consulted and modified during the following stages. These structures are automatically passed to other functions inside the CMD software.

2.5.3 Function 2: Model structure configuration

Next, the user can select one or more cells of which the model structures should be identified and calibrated. The tool goes in a step-wise manner through the set-up procedure of each module and proposes a parameter set for the elected model structure. The structures for each module are outlined in §2.3 of this chapter. The algorithms or toolboxes required to configure each module are incorporated in the CMD software itself and are listed in Table 2.2. Two different methods are included to estimate TFs, namely the RIV algorithm (see §2.3.1.4) and a function from MATLAB's System Identification Toolbox. The latter function allows constraining parameters during estimation. After calibration of a module, its simulated values are plotted together with the calibration and validation data. The user can choose to adopt the proposed model structures and parameter sets, or select new ones. These objective criteria depend on the selected model structure and are discussed in §2.3. Hence, although the tool is highly automated, expert knowledge can still be incorporated. Lastly, ANNs are transformed before being stored again in organized MATLAB data structures. These ANNs are converted into one or two dimensional lookup tables with variable resolution (see also Wolfs and Willems, 2014). Such low dimensional tables can be visualized easily, thereby allowing the user to assess the extrapolation behaviour and the physical soundness of the ANN. Furthermore, this transformation facilitates their incorporation in the final calculation scheme written in C and these tables are computationally less demanding to estimate the output.

2.5.4 Function 3: Model assembly

The third main function of the CMD software is the creation of the script that represents the configured conceptual model and the required boundary files to run simulations. This is done by assembling all configured modules for selected reservoirs. Since the user specified the river topology, this interfacing between different modules and variables is completely automated. The boundary data required for the selected events that will be simulated is prepared, together with arrays to store the states of the model that will be calculated. The latter allows for memory pre-allocation, which speeds up the simulation. It is vital to be able to include only parts of the river in a model, so the influence of one reservoir on another and vice versa can be studied, and modules requiring adjustments can be identified easily.

Table 2.2: Algorithms and toolboxes included in the CMD software to identify and calibrate model structures.

Model structure and functionality	Module	Algorithm and/or toolbox
Weir equation optimization	Q	MATLAB implementation of SCE-UA (Duan et al., 1992) MATLAB Optimization Toolbox
Transfer Function	Q	RIV – CAPTAIN Toolbox (Young et al., 2007) MATLAB System Identification Toolbox
Non-Linear Transformation	Q	SDP – CAPTAIN Toolbox (Young et al., 2007)
Artificial Neural Network	Q/WL	MATLAB Neural Network Toolbox
Piecewise Linear Relationship	Q/WL	MATLAB algorithm PWL LSQ fit
M5' Model Trees	WL	MATLAB toolbox by Jekabsons (2010)

2.6 Case study

2.6.1 Study area and events

A conceptual model was built of the Marke River in Belgium to demonstrate the presented modelling methodology and software tool. The developed conceptual model was afterwards used for the impact analysis of flood retention basins along the Marke River on river floods in the Dender catchment. The results of this study are described in Chapter 6.

The Marke is a non-navigable and flood-prone river with numerous floodplains, some of which are embanked. The flow is controlled by 11 gated weirs that rely on a set of logic rules and water level measurements to determine their gate positions. Most of these gated weirs are equipped with parallel fishways to allow fish migration. The aim is to create a model that predicts the behaviour of the considered river during normal and high flow conditions and requires a minimal computation time. A small part of the Dender River which receives water from the Marke was also incorporated in the conceptual model. Therefore, interactions at the confluence of both can be studied.

A large amount of flow and stage measurements is required to set up an accurate conceptual model due to the grey box nature of the modelling methodology. In addition, precise gate level readings are vital, since these can significantly affect the flow over gated weirs in the river. These extensive data requirements are often a key confining factor when developing simplified models (see also Hunter et al., 2007). Because reliable field data were only available at a few locations for this study, "virtual" data were employed instead. These data were generated by simulating different events with a detailed full hydrodynamic IWRS model of the Marke River. The detailed model incorporates cross-sections measured every 50 meters and includes all important hydraulic structures. It was calibrated by the Flemish Environment Agency (VMM) using gauge measurements, field surveys, maps of recently flooded areas and digital elevation models. Different PDM models (Moore, 2007) distribute the rainfall runoff to the river. The logic rules that determine the gate positions are also implemented in the IWRS model. During simulations, calculated water levels and previous gate positions are used instead of measured stages in the logic rules. Note that this IWRS model was also used in Chapter 5 to model rating curves affected by hysteretic behaviour. A more comprehensive discussion of a looped water level-flow relationship near a movable weir can be found in §5.3.1.

Simulation results of the flood events of November 2010 and March 2008 were used for calibration of the conceptual model, while the event of January 2011 acts as validation data. The storm of November 2010 was very extreme and caused significant damage in the entire region. The March 2008 storm is characterized by limited flooding, while the January 2011 storm caused serious inundations at multiple locations. During the November 2010 storm, the highest water level was measured since the installation of fixed limnigraphs in 1976. Figure 2.7 shows the measured water levels near the city of Viane, which is a few kilometers upstream of the confluence of the Marke and Dender Rivers, together with predicted values by the InfoWorks RS and the derived conceptual model (see §2.6.2).

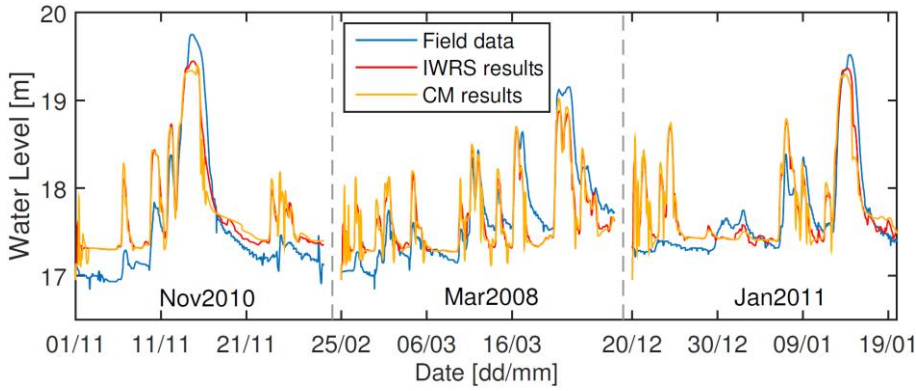


Figure 2.7: Water level measurements near Viane, and InfoWorks RS and conceptual model (CM) simulation results for the events of November 2010 ('Nov2010'), March 2008 ('Mar2008') and January 2011 ('Jan2011').

Significant water level differences can be noticed between the IWRS simulated water level and available field data, especially for the November 2010 storm (Figure 2.7). This can, however, be expected due to the very extreme character of this event. Many of the measurement devices malfunctioned during this storm, due to debris or simply because they themselves were flooded. In such situations, the gated weirs are configured to open completely, thereby reaching their maximum capacity. This behaviour was not captured by the IWRS model, since it relies on simulated water levels in its gate position determination. In addition, other factors may explain the deviations, such as parameter and model structure uncertainties of both the IWRS model and the used hydrological PDM models, and measurement errors of rainfall and the investigated water level series. Even though the IWRS model cannot simulate the river's state very precisely, it can be used perfectly to exemplify the developed conceptual modelling approach. Figure 2.7 already presents the simulated water levels at Viane of the conceptual model, although the set-up of this model is only discussed in the following section. It shows that the additional uncertainty on water level predictions induced by the simplifications of the conceptual model is very limited compared to the mismatch between the detailed IWRS model and the field data. Note that the conceptual model does not aim to emulate reality more precisely than the detailed model. In fact, due to the data-driven character of the conceptual models, they cannot outperform the hydrodynamic models to which they were calibrated (naturally, improved predictions are sporadically possible due to variations in the models and inputs though). The conceptual models can only mimic the dynamics that are present in the simulation results of the detailed models.

Three synthetic events, all variants of the November 2010 storm, with the following characteristics were created for additional validation (and to assess the extrapolation behaviour, see also §2.6.3.2):

- Nov2010 – S1: All incoming flows are increased by 20%;
- Nov2010 – S2: The November 2010 storm is simulated twice after each other;
- Nov2010 – S3: The width of the peak of all incoming flows is halved, so the rising and falling limbs of the incoming hydrographs are much steeper.

The gate positions of hydraulic structures were determined using the implemented logic control rules of the IWRS model during simulations. These storms allow assessing the extrapolation capabilities of the conceptual model. The Nov2010-S2 storm can be used to investigate the emptying of floodplains better. In addition, this event is more extreme than the historical November 2010 since not all floodplains are empty when the second peak arrives. The fast rising and decaying incoming flows of the Nov2010-S3 storm could induce additional dynamic effects, and test the response of the gate level logic controls.

2.6.2 Conceptual model configuration

To set-up the conceptual model, the workflow discussed earlier is followed (see Figure 2.1). First, all data is gathered that is required to configure the model: the river network and logic controls of movable structures are exported from the IWRS model to csv-files, together with the IWRS simulation results of flows, volumes, water levels and gate positions. The IWRS simulation results of each event used for calibration and validation comprise 30 days and were saved every 100 seconds. The conceptual model employs the same rainfall-runoff as in the detailed hydrodynamic model.

2.6.2.1 Topology conceptualization

The next step of the model set-up is the division of the river and accompanying floodplains into an arrangement of storage cells. River segments confined by controllable hydraulic structures or sections that influence the flow in the river significantly are merged into one storage cell in the conceptual model. The floodplains are conceptualized likewise, although they are most often bounded by dikes instead of hydraulic structures. All interconnected storage cells of one

floodplain are grouped into a "floodplain cluster" (FPC). Hence, such FPC can contain multiple distinct storage cells and represents one floodplain. Figure 2.8 shows the river topologies of the detailed IWRS and conceptual models.

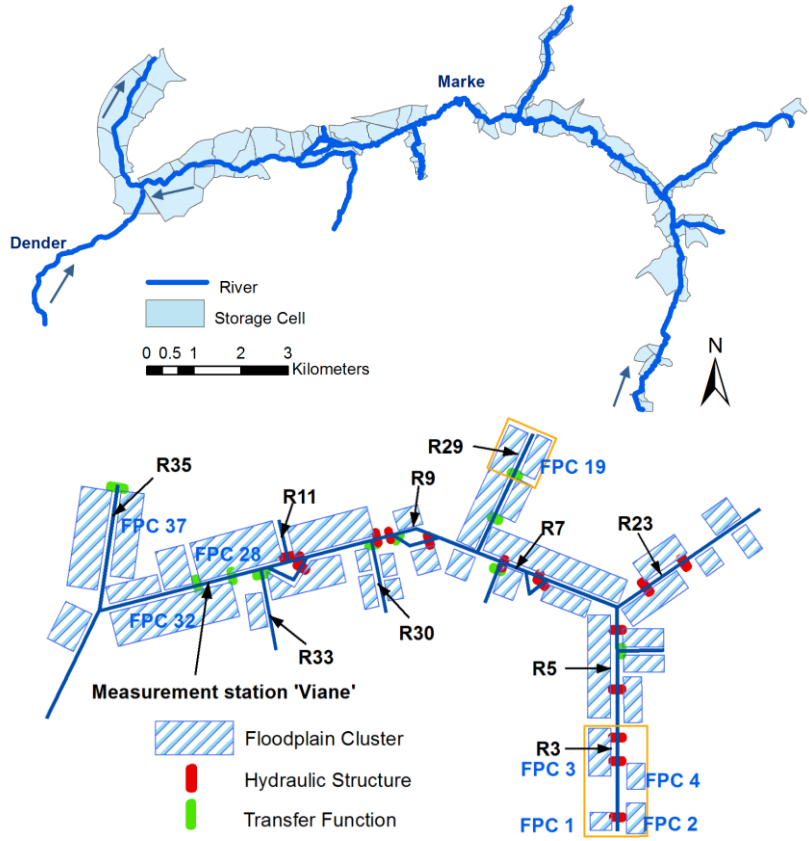


Figure 2.8: River network in InfoWorks RS used to set up the conceptual model (top) and in the conceptual model (bottom). The set-up of the cells enclosed by the orange rectangles is shown in more detail in Figure 2.9 and discussed in Sections 2.6.2.2 and 2.6.2.3).

Next, the templates defining the network topology are imported in the CMD software and the required data is extracted from the collected data sets by the tool. Next, different model structures are identified and calibrated using the CMD software. Two sections of the conceptual model are discussed more in-depth to illustrate some of the available modules and the calibration process. These two sections are marked by the orange rectangles in Figure 2.8, and are shown more

closely in Figure 2.9. The other elements of the conceptual model are configured similarly.

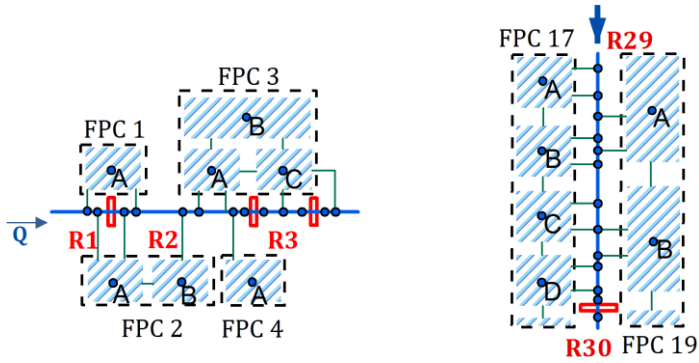


Figure 2.9: Schematic conceptual network of (left) three river reservoirs ('R1' to 'R3') and four floodplain clusters ('FPC 1' to 'FPC 4'), and (right) two river reservoirs ('R29' and 'R30') and two floodplain clusters ('FPC 17' and 'FPC 19'). Water levels are calculated at each blue node. The discharge is estimated through the thin green lines between different cells. Red rectangles represent hydraulic structures.

2.6.2.2 Area 1: Retention basin Herne and surrounding cells

The first area focuses on the most upstream part of the river and comprises several floodplains and three controllable hydraulic structures that regulate the flow in the river. These structures demarcate the first three storage cells (see "R1" to "R3" in Figure 2.9 - left). It is essential to include them explicitly in the conceptual model, since they significantly affect the flow and consequently, their control rules need to be translated to the conceptual model as well. It is possible to adopt the parameters of these weirs directly from the detailed model, because the conceptual model employs the same equations (see §2.3.1.1 for more info on this module). Alternatively, it is possible to optimize the parameters for given stage and discharge time series with one of the optimization algorithms incorporated in the CMD, or use self-defined discharge relationships. The logical rules (PLC) that determine the gate openings of the hydraulic structures are transferred from the InfoWorks RS model to the conceptual model as well (see §2.3.4 and §2.5.2). Next, the locations at which the water level should be calculated are identified based on the defined topology, such as up- and downstream of weirs, near floodplains embankments or for PLC. The CMD software proposes simple piecewise linear hypsometric curves (see §2.3.3.1) for each of these locations (see e.g. Figure 2.10c), relating the water level to the

storage of the cell. Because the hypsometric curves are almost univocal, it was not necessary to use more advanced model structures or include a state dependent correction (see §2.3.3.2). Note that it is impossible to use rating curves (which presume a relationship between the discharge and the water level) due to the presence of controllable structures. The final module closes the water balance by summing up all incoming and outgoing flows. Naturally, this module does not require further calibration, since the relevant flows are deduced from the defined topography.

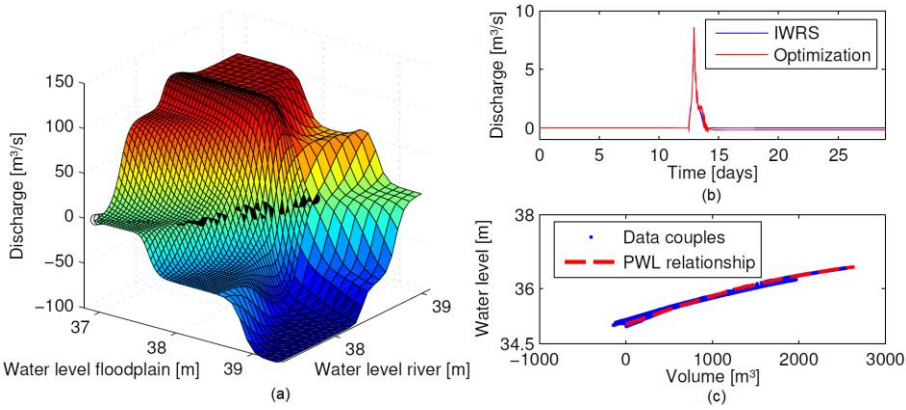


Figure 2.10: Grid with variable resolution that emulates the outcome of the trained ANN and is used to estimate the flow (left), approximated flow in function of time by the ANN for the calibration event (top right) and calibrated PWL hypsometric curve of a water level in cell "R3" (lower right).

Floodplain cluster 3 is partially used as retention basin during high flood events. It consists of three storage cells that are filled and emptied via fixed embankments and flapped orifices. Hence, filling and emptying is controlled by regulating the water level in the river. Present orifices are modelled separately by adopting the relationships and parameters from the detailed model directly. The flow over dikes is modelled by several spills in the IWRS model that take the irregular form of the embankments into account. These overflows are lumped to limit the number of calculation nodes in the conceptual model and are emulated by fixed broad crested weir equations. The CMD software considers different combinations of up- and downstream water level locations for optimizing the parameters. First, the dike level is estimated for a set of up- and downstream water level time series by identifying the first overtopping of the dike (i.e. flow different from zero). Next, the parameters of the broad crested weirs are

optimized via the interior-point algorithm incorporated in the CMD software. The combination of water levels yielding the highest accuracy is retained by the software, together with the calibrated parameter sets. The water levels and volume calculations are similar as in the cells described above.

Artificial neural networks (ANNs) are employed to estimate the flow from and to the storage cells of FPC2 to illustrate their use and configuration procedure (see §2.3.1.2). Note that due to the rather simple dynamics exhibited by the investigated flows, the use of weir equations showed comparable emulation performance in this case. The use of ANNs is particularly beneficial when flows are lumped on much larger scales than in this study (see e.g. Wolfs and Willems, 2014a), or when the embankments have very irregular shapes. Each flux is represented by a separate ANN, which uses the up- and downstream water levels of the flux as inputs. After the CMD software has estimated the dike level for a set of up- and downstream water levels, the training data (consisting of the flow and these water levels) are pre-processed by removing all samples with both water levels below the dike level. The training of the ANNs is purely focused on non-zero flows, while a simple binary classifier (if/else-condition) determines if the water levels overtop the dike. During simulations, the outcome of ANNs in the future conceptual model is only assessed when at least the up- or downstream water level surpasses the dike level, while a flow of zero is returned for lower water levels. This not only reduces calculation time, it also ensures that the dike level determines the moment of overtopping. Next, different neural networks with a limited amount of hidden nodes are trained in the CMD software using the early stopping technique. As done for the weir equations, this process is repeated for multiple combinations of up- and downstream water levels. Again, the set yielding the best performance is preserved. This set of water level locations is used to retrain the neural network with different numbers of hidden nodes. Trial and error showed that the use of ANNs with eight hidden nodes yields the best emulation results. Finally, the CMD tool transforms the ANN into a two dimensional lookup table (see also Section 2.5.3 for further details). Figure 2.10 shows the obtained neural networks of the flux between storage cells A and B (see Figure 2.9– FPC2). The automatically generated grid representing the ANN has a variable resolution: subspaces that are more likely to be addressed during a simulation or show more distinct variations have a denser resolution than other spaces. This speeds up the calculation during simulations since it reduces the grid size and hence the look-up operations. Secondly, the form of the grid is physically meaningful (flow is defined positive from cell A to B) and smooth. However,

ANNs remain data-driven and can therefore only be used effectively in a limited domain. To limit extrapolation errors, the boundaries of the grid are extended horizontally.

2.6.2.3 Area 2: Cell 'R29'

The second discussed area contains a culvert and several floodplains. These floodplains are again modelled using weir equations in combination with hypsometric curves to estimate the water levels. The culvert is located in the river itself and limits and delays the flow. Hence, such structure can cause backwater effects, but analysis of the available flow and stage data shows that these effects have a recurrent character. The culvert cannot be controlled externally and can therefore likely be emulated by a transfer function (see §2.3.1.4). Since the flow towards the adjacent floodplains is modelled explicitly, the incoming flow of the transfer function is equal to the upstream flow minus the discharge of the floodplain overflows. Alternatively, it is possible to perform a non-linear transformation of the upstream flow (see §2.3.1.4). Identification and estimation using the built-in functions from the System Identification Toolbox incorporated in the CMD tool yields a simple first order, continuous transfer function. The steady state gain of the TF is constrained to unity to preclude the loss or creation of water during simulations, and certifies that the water balance is closed. To speed up the calculation and simplify the conceptual calculation scheme, the obtained transfer function is converted to the discrete domain for different sampling times. Finally, rating curves are determined that relate the water levels required for the calculation of the flow towards floodplains to the discharge estimated by the transfer function.

2.6.3 Results

2.6.3.1 Accuracy assessment

All configured modules are assembled into one model script by the CMD tool. The derived conceptual model was then used to simulate all six investigated events. The complete model contains 144 storage cells, whereby at 331 and 316 different locations respectively water levels and discharges are estimated. Only a few crucial and important calculation nodes are selected spread over the model for assessing the accuracy, since it is impossible to present the performance at all locations clearly. The conventional Nash-Sutcliffe efficiency (NSE) is calculated for each location in the selection list:

$$NSE = 1 - \frac{\sum_{i=1}^{n}(Y_{"observed",i} - Y_{predicted,i})^2}{\sum_{i=1}^{n}(Y_{"observed",i} - \bar{Y}_{"observed"})^2} \quad (2.7)$$

with *Y* the investigated variable, *i* the value at a specific time step, *n* the number of entries of the time series. Predicted refers to simulation results of the conceptual model. The simulation results of the InfoWorks RS model used to configure and validate the conceptual model are denoted as "observed" data.The NSE can range from -∞ to 1. An efficiency of unity implies a perfect match between observed and predicted data, while a value of zero indicates that the predicted values are as accurate as the mean of the observed data. A negative NSE-value occurs when the observed mean is a better predictor than the model being used. Hence, the closer the efficiency is to 1, the more accurate the model is. Not enough data from gauges was available to perform a thorough comparison with in-situ measurements.

Table 2.3: NSE values of the total outgoing flows for selected storage cells (in the river) for the calibration ('Nov 2010' and 'Mar 2008') and validation ('Jan 2011' and 'Nov 2010 – S1/S2/S3') events. These values compare the simulation results of the conceptual and InfoWorks RS models.

	Nov 2010	**Mar 2008**	**Jan 2011**	**Nov 2010 – S1**	**Nov 2010 – S2**	**Nov 2010 – S3**
R3	0.997	0.996	0.992	0.996	0.995	0.991
R5	0.974	0.900	0.922	0.972	0.974	0.936
R7	0.958	0.896	0.923	0.966	0.960	0.928
R9	0.965	0.870	0.927	0.950	0.968	0.915
R11	0.979	0.926	0.954	0.965	0.982	0.957
R23	0.995	0.998	0.996	0.993	0.996	0.994
R29	1.000	1.000	1.000	1.000	1.000	1.000
R30	0.983	0.998	0.986	0.892	0.987	0.980
R33	0.964	0.734	0.958	0.936	0.971	0.945
R35	0.985	0.977	0.981	0.982	0.988	0.979

Table 2.4: NSE values of the water level predictions for selected storage cells (in the river) for the calibration ('Nov 2010' and 'Mar 2008') and validation ('Jan 2011' and 'Nov 2010 – S1/S2/S3') events. These values compare the simulation results of the conceptual and InfoWorks RS models.

	Nov 2010	**Mar 2008**	**Jan 2011**	**Nov 2010 – S1**	**Nov 2010 – S2**	**Nov 2010 – S3**
R3	0.977	0.883	0.953	0.981	0.972	0.944
R5	0.895	-1.248	0.513	0.913	0.916	0.768
R7	0.569	0.085	0.700	0.569	0.734	0.388
R9	0.965	0.818	0.916	0.953	0.967	0.897
R11	0.969	0.702	0.917	0.974	0.964	0.917
R23	0.967	0.972	0.958	0.938	0.976	0.980
R29	0.999	0.999	0.999	1.000	1.000	0.999
R30	0.996	0.996	0.995	0.996	0.997	0.992
R33	0.942	0.659	0.918	0.940	0.959	0.892
R35	1.000	0.780	1.000	1.000	1.000	1.000

Table 2.3 shows the calculated NSE values for the total outgoing flow (i.e., the sum of all flows defined positively from the examined storage cell to others) and Table 2.4 shows the NSE-values for the water level predictions at the selected locations. Although a storage cell can contain multiple nodes where the water level is calculated, only one water level is used here in the assessment since those water levels are usually highly correlated anyway. The NSE values for all outgoing flows are very close to unity, indicating a good fit. The lower NSE values for the water levels indicate that the predicted states of the conceptual model deviate a little from the simulated water levels by the detailed model. The water level prediction of reservoir 'R5' even has a negative NSE value, which implies that the mean of the time series provides a better fit than the model. The lower performances for the stage predictions in reservoirs 'R5' and 'R7' are caused by the gated weirs that aim to regulate the water levels in a very narrow band as illustrated by Figure 2.11d. The differences in water level predictions remain in these two reservoirs very limited in absolute terms, but result in poor NSE values. The conceptual model manages to emulate the InfoWorks RS results in reservoir 'R29' almost perfectly, as shown by the NSE values of nearly unity. This reservoir is located in a tributary of the Marke that is not affected by backwater effects or controllable hydraulic structures, and exhibits simple flow dynamics. Note that

the results of some validation events outperform those of calibration events. This is not surprising due to the mechanistic nature of the modelling approach: each module is configured individually and based on physical principles. In addition, the dynamics encountered during some calibration events might be more complex than during certain validation events, leading to poorer results.

Table 2.5: NSE values of the water level predictions for selected storage cells (in floodplains) for the calibration ('Nov 2010' and 'Mar 2008') and validation ('Jan 2011' and 'Nov 2010 – S1/S2/S3') events. These values compare the simulation results of the conceptual and IWRS models. During the 'Mar 2008' storm, no flooding occurred in these floodplains in both the conceptual and IWRS models and are therefore marked as 'n.a.'.

	Nov 2010	**Mar 2008**	**Jan 2011**	**Nov 2010 – S1**	**Nov 2010 – S2**	**Nov 2010 – S3**
FPC 36 – cell C	0.996	n.a.	0.986	0.982	0.997	0.995
FPC 28 – cell I	0.977	n.a.	0.926	0.984	0.874	0.845
FPC 37 – cell E	0.980	n.a.	0.996	0.973	0.998	0.982
FPC 28 – cell H	0.983	n.a.	0.728	0.989	0.749	0.757
FPC 32 – cell D	0.966	n.a.	0.974	0.960	0.962	0.930
FPC 36 – cell A	0.926	n.a.	0.743	0.876	0.979	0.978
FPC 37 – cell B	0.944	n.a.	0.972	0.934	0.682	0.878
FPC 37 – cell D	0.964	n.a.	0.984	0.956	0.887	0.917
FPC 28 – cell G	0.909	n.a.	0.430	0.860	0.973	0.495
FPC 36 – cell C	0.995	n.a.	0.987	0.986	0.997	0.996

Next, the 10 floodplain cells with largest storage during the "Nov2010" event are selected, together with the largest storage cell from Floodplain cluster 3 which configuration was discussed (see §2.6.2.2). Table 2.5 summarizes the NSE-values for the calculated water levels in each storage cell. Again, NSE-values close to unity can be found, except for some locations during the "Jan2011" storm. During this event, some water levels in the river fluctuate around the dike levels of certain floodplains. Small inaccuracies in such river water level predictions make the difference between overtopping of the dikes and hence filling the floodplains, or remaining just under the dike level. This can easily lead to significant differences in water level predictions in floodplains, and hence lower performance criteria. Figure 2.11c shows the maximal water depths in all floodplain storage cells during the simulated events. It is clear that the storage

cells that have the greatest maximal water depths during simulations with the IWRS model are emulated very accurately by the conceptual model, but some scatter is noticeable for floodplain cells with smaller inundation depths.

Finally, the water balance is checked explicitly to ensure no water is lost or created during simulations. Comparison of the volume water stored in the river and floodplains with the incoming and outgoing flows shows that there is only very little difference. The volume error did not exceed 10^{-6} m³ and is hence caused by rounding errors.

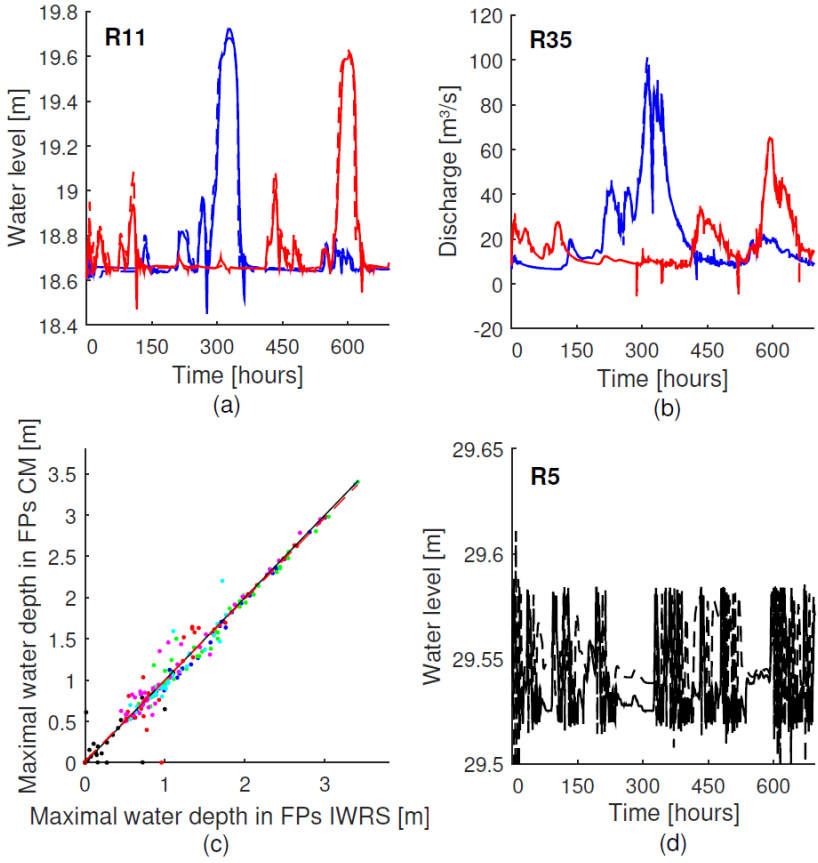


Figure 2.11 (first part): Water level and discharge simulation results of the detailed InfoWorks RS (dashed line) and conceptual (CM, full line) models.

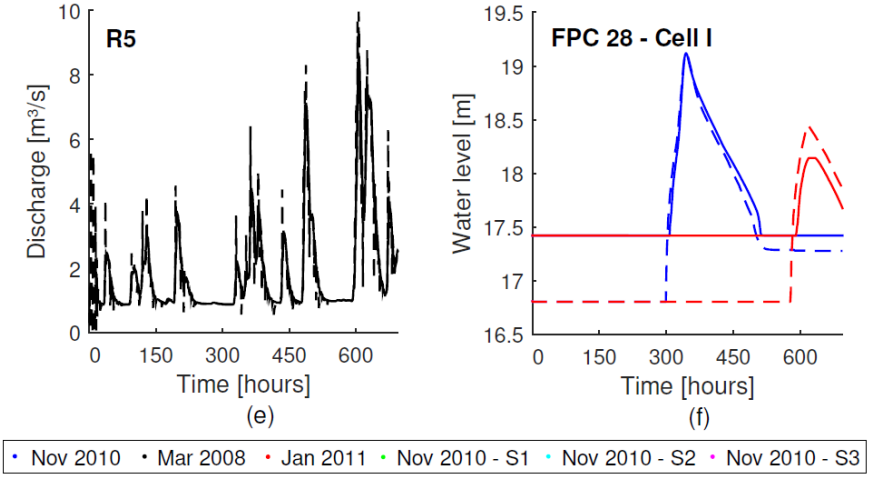


Figure 2.11 (continued): Water level and discharge simulation results of the detailed InfoWorks RS (dashed line) and conceptual (CM, full line) models.

2.6.3.2 Extrapolation performance

Data-driven models often have difficulties with extrapolation. Since the presented approach reasons from a mechanistic, physically based concept (i.e. the network topology as well as most incorporated model structures and parameters can be interpreted in physically meaningful terms), it can be applied for extrapolation to some extent though. To assess the extrapolation behaviour of the configured conceptual model and the approach in general, its different modules and how each will respond during extrapolation are briefly discussed. Volumes and gate levels are simply predicted using the continuity equation in its discrete form and programmable logic controls respectively, and are hence not affected when values are simulated beyond the calibration range. Water levels are simulated via piecewise linear relationships or M5' model trees which depend on a flow or volume (see §2.3.3.1). During model set-up, one can easily estimate the outcome of these structures for larger inputs. Finally, different structures are incorporated to simulate flows. Hydraulic structure equations (see §2.3.1.1) can be used for extrapolation, since they are physically based and effects such as drowned flow are accounted for (if present during calibration). To estimate the extrapolation behaviour of ANNs (see §2.3.1.2), the network is transformed in a one- or two-dimensional lookup table that is visualized during the calibration process. This allows the user to see the outcome of ANNs for data other than the training set. Transfer functions (see §2.3.1.4) act similarly to an arrangement of linear reservoirs, which are probably familiar to many readers.

Hence, the modeller can (to a great extent) predict *how* the model will behave during extrapolation, but not how *accurate* it will be. Indeed, due to the data-driven character of the approach, only effects that are already incorporated in the calibration data set can be captured. Hence, if the river exhibits fundamentally different behaviour (e.g. overtopping of dikes that were not flooded in the calibration set), the model cannot account for such effects. An advantage of the presented approach is that it allows for partial recalibration specifically targeted at emulating this new behaviour, while other model structures are not affected.

Table 2.3 and Table 2.4 show that the conceptual model emulates the simulated flows in the IWRS model very accurately for the three extrapolation events (Nov 2010 – S1 to S3). All NSE values, except for one, surpass 90%. The accuracy of the water level predictions in the river and floodplains (see Table 2.4 and Table 2.5) does not drop significantly either, but there is some additional scatter noticeable for the maximal attained water depths in floodplains (see Figure 2.11c). Higher flood depths are still modelled very accurately though. The emulation of the water depth in FPC37 – cell B dropped significantly for event 'Nov 2010 – S2'. These deviations are caused by the steepness of the defined hypsometric curve of this cell: minor deviations in low volume predictions in floodplain cells result in large water level errors (see Figure 2.12). This very steep hypsometric curve can be caused by topographic features or by filling Preissmann Slot-type elements in the detailed IWRS model. For the same reason, some water level predictions in FPC28 for the 'Nov 2010 – S3' event are inferior compared to those of the historical events.

In general, one can conclude that the emulation capabilities of the configured conceptual model are adequate. Although the synthetic storms differ significantly from the calibration events, water levels and floodplains are still modelled satisfactorily. The overall high NSE-values for water levels and flow predictions at various locations in the catchment indicate that the conceptual model succeeds in simulating the river's and floodplains' states accurately, which was its main performance goal.

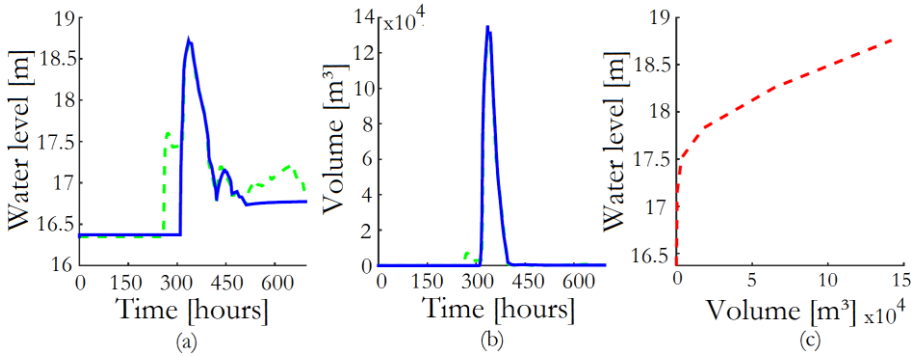


Figure 2.12: Water level (a) and volume (b) predictions for FPC37 – cell B for the 'Nov 2010 – S2' event of the IWRS (green, dashed line) and conceptual (blue line) models. Although the water level predictions of the conceptual model deviate significantly from the IWRS simulations, the volume predictions are very accurate. These large water level deviations are caused by the steep hypsometric curve of this cell at smaller volume values (c).

2.6.3.3 Computational efficiency

Some reservoirs have a limited storage capacity, while their water level can vary rapidly. Such reservoirs are prone to instabilities and can cause fluctuations that spread rapidly over the entire model. To overcome this, the discrete variable time step solver presented in §2.4 is applied to 11 strategically selected storage cells with limited capacity. If the non-linearity estimator of such storage cell surpasses a specified threshold (set to 1.9 in this case) and hence oscillations originate, the time step is reduced to 12.5 seconds and the calculation of this storage cell is repeated. Storage reservoirs without variable step solver are calculated at the fixed interval of 100 seconds.

Simulating the 25000 time steps of the "Nov2010" event with the conceptual model with a fixed time step takes 0.83 seconds (averaged over 100 runs) on a computer with i7 processor clocked at 3.40 GHz and 16 GB RAM memory using a single core, while the calculation time amounts 1.12 seconds if the variable time step solver is used. This is much faster compared to the approximately 50 minutes required to do the same simulation with the detailed IWRS model. Figure 2.13 illustrates the effectiveness of the adaptive time step scheme (see §2.4). The estimated volumes in the conceptual model start to oscillate rapidly when a fixed time step is used for the entire simulation (i.e. without adaptive time step), while such fluctuations are avoided almost completely when the variable time step

scheme is applied for instability-prone storage cells. The time step is reduced and the solution is iterated when the stability indicator shown on Figure 2.13 is equal to one. After the peak of the storm, the variable time step solver is still activated in FPC2 – Cell B, because the volumes in both the IWRS and conceptual models balance just above zero and are in unstable equilibriums.

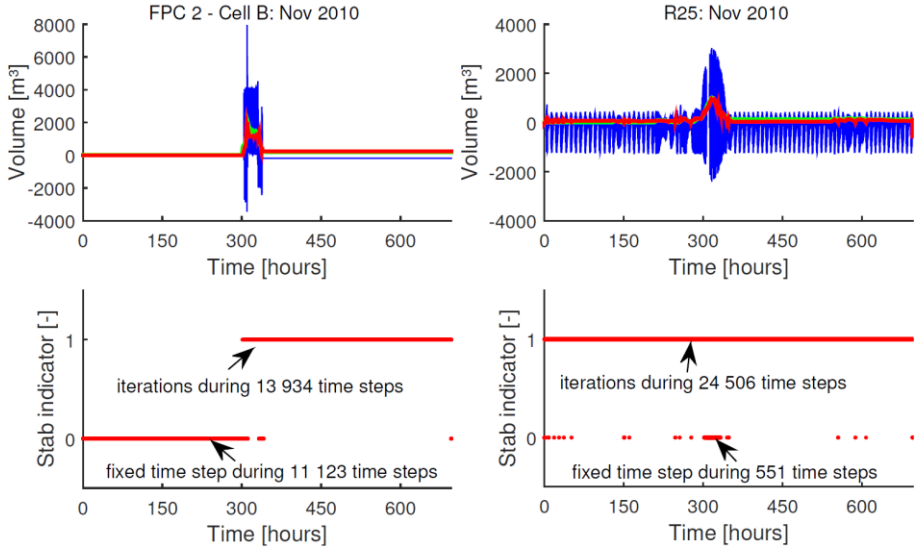


Figure 2.13: Simulation results of the storage in two storage cells for the "Nov2010" event of the IWRS (green) and conceptual models with (red) and without (blue) adaptive time step.

2.7 Conclusions

This chapter discussed the development of a novel and highly flexible conceptual modelling methodology for river hydraulic computations that can cope with the complex dynamics of rivers and floodplains. The approach combines advanced model structures and matching optimization techniques in a mechanistic and modular framework. The methodology is based on the storage cell concept and supports high level conceptualization of river segments, but it is also possible to incorporate essential elements of the river system explicitly, such as dike levels, controllable weirs, pumps and sluices. Such modular design ensures that the river dynamics can be mimicked accurately at reach scale, and the models are suitable for a wide variety of applications.

To facilitate and speed up the set-up of these conceptual models, a comprehensive and semi-automatic tool (Conceptual Model Developer, CMD) was developed that assists the modeller during model configuration. Close integration was foreseen with the InfoWorks RS software, but it can also handle spreadsheet data sets and is therefore compatible with virtually all software packages.

Alongside this modelling approach and software tool, a new calculation scheme was developed that supports a time and space varying time step. This maximizes computational efficiency while avoiding instabilities. The proposed scheme has great potential if some parts of the model are more prone to instabilities than others, and can also be applied to other storage cell models, including the sewer modelling approach presented in Chapter 3.

To illustrate the applicability and efficiency of the modelling approach and solution scheme, a conceptual model was created of the Marke River in Belgium. The results show that the conceptual model manages to emulate a detailed full hydrodynamic InfoWorks RS model accurately for all investigated events, while the calculation time is reduced by more than 2000 times compared to a full hydrodynamic model.

3

Conceptual modelling approach and software for sewer systems

3.1 Introduction

This chapter presents the newly developed approach to conceptually model sewer systems. The methodology emanates from the same pragmatic perspective and modelling principles as the modelling approach for rivers presented in Chapter 2, but is tailored to urban drainage systems. Therefore, an adapted model topology and other model structures are employed. Again, a software tool was developed that incorporates the conceptual modelling approach. This tool was implemented mainly independently from the software to conceptually model river systems, allowing easier maintenance, expansion and future implementation in other systems. Naturally, the tool for conceptual sewer modelling shares certain fundamental programming scripts and functions with the tool developed for conceptual river modelling. A case study demonstrates the modelling approach.

This chapter is structured similarly to Chapter 2 to facilitate comparison between both modelling approaches. First, the conceptualization of the sewer system is discussed. Section 3.3 presents the different model components. Compared to the modelling approach for rivers, there are two main differences with the methodology for sewer systems. The latter only calculates volumes and flows, and

hence does not simulate water levels or gate levels. Secondly, the sewer inflow (consisting of rainfall runoff, dry weather flow, trade and industrial flow, etc.) is also modelled in this approach, since the modelling of inflows is highly intertwined with the conceptual model topology. Detailed information on parcel level is often necessary to estimate the incoming flow. Next, Section 3.4 describes the development and functionality of the software to semi-automatically set up the conceptual sewer models. As for the software for river systems, this tool is also named CMD. Again, an interfacing is foreseen with the InfoWorks software. Lastly, Section 3.5 presents a case study to illustrate the system conceptualization and the use of model structures. A conceptual model of part of the city of Ghent is configured from simulation results of an InfoWorks CS model.

The chapter is based on the following publications:

> WOLFS, V., VILLAZON GOMEZ, M., WILLEMS, P. (2013). DEVELOPMENT OF A SEMI-AUTOMATED MODEL IDENTIFICATION AND CALIBRATION TOOL FOR CONCEPTUAL MODELLING OF SEWER SYSTEMS. WATER SCIENCE AND TECHNOLOGY, 68 (1), 167-175.
>
> WOLFS, V., WILLEMS, P., SUBMITTED. MODULAR CONCEPTUAL MODELLING APPROACH AND SOFTWARE FOR SEWER HYDRAULIC SIMULATIONS. ADVANCES IN WATER RESOURCES.

3.2 Conceptualization of the sewer system

3.2.1 Modelling principles

The presented conceptual modelling approach reasons from the same principles as discussed extensively in §2.2.1 and are therefore not repeated in detail. Again, a data-driven mechanistic and modular framework is adopted, as was already suggested by several authors in the past (Schütze et al., 2004; Vanrolleghem et al., 2005; Salvadore et al., 2015). The use of such an approach is advocated, since it seems impossible to find one general set of equations that is able to describe the flow accurately in all sewer systems under varying conditions, and can be solved efficiently at the same time. Sewer systems can respond very differently to inputs and conditions, such as when backwater effects prevail (Sartor, 1999). The conceptual model topology is also based on the storage cell concept, allowing the modeller to lump processes on different spatial scales and to solely focus on the prevalent processes for the intended applications.

The modelling approach for sewer systems differs on several other aspects from the methodology for rivers. Naturally, a modified model topology and alternative model structures are incorporated in this approach, as is discussed in the next sections. Each incorporated predefined set of equations is tailored to tackle specific dynamics encountered in sewer systems. The relationships do not hinge on stringent a priori assumptions that limit their usability. This results in very flexible relationships that can be employed at various scales and under different conditions. Secondly, as in most hydrological models, the approach only predicts volumes and flows. Simulating water levels in sewer systems is not considered due to its complicatedness, and because water levels are often not required in applications. Furthermore, experiments show that simulating water levels is not pivotal for achieving accurate flow estimations. As demonstrated in this chapter (see §3.5), excluding water level simulations does not preclude accurate CSO predictions and volume estimations, even when the flow is influenced by pumps or other hydraulic structures (such as valves, weirs and sluices). In fact, many studies involving model-based predictive control of such structures use models that do not incorporate water level predictions (e.g. Schütze et al., 2004; Vanrolleghem et al., 2005; Vezzaro et al., 2014). Therefore, the presented modelling methodology only simulates volumes and flows. Water levels can be used as boundary conditions though, which is for instance often crucial to model the interaction with rivers accurately. Thirdly, the inflow into the sewer system is also conceptualized and incorporated in the modelling approach and CMD software. To determine the total inflow, many different parameters and time series on parcel level must be gathered and combined. The conceptual model topology, and hence also the lumping of the information of parcels, is determined in an iterative procedure (see also §3.2.2). To ensure that the conceptual model topology can be modified quickly, the modelling of inflows is also included in the methodology and software. The tool automatically assembles all required data to calculate new inflow time series when the model topology changes.

3.2.2 Model configuration procedure

Setting up conceptual models according to the presented methodology is highly similar to the procedure discussed in §2.2.2. Since the modelling approach for urban drainage systems differs on some aspects from the developed approach for rivers, the procedure is briefly outlined, with special attention to the differences. The configuration process comprises several main steps as depicted in Figure 3.1: collecting the required data, defining the conceptual model topology, identifying

and calibrating model structures that emulate the flows in the system, assemble all configured elements in a model script, and finally simulate events and assess the accuracy.

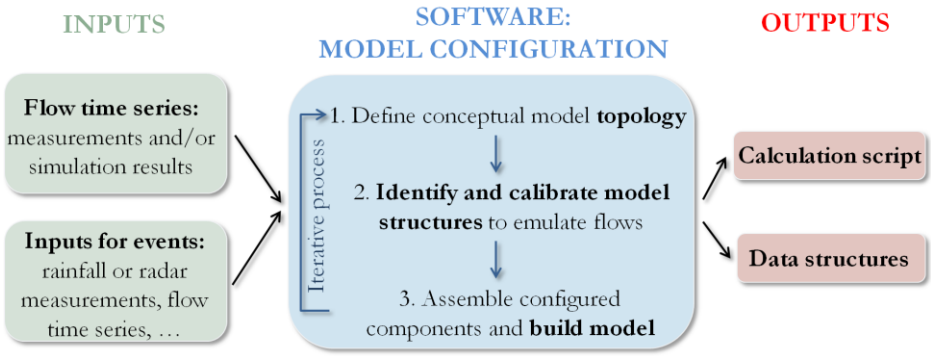


Figure 3.1: Schematic overview of the set-up procedure of conceptual models and the functionality of the software tool.

First, the required data must be gathered to configure the conceptual model. Since the presented methodology is generally driven by flows and volumes, acquiring accurate and reliable time series of flows is an essential element of the set-up process. Time series of water levels, such as stages in rivers or ponds, can optionally be employed as boundary data during model configuration and simulation, but will not be estimated in the presented modelling methodology. If a rainfall-runoff component will also be included in the conceptual model, time series of rainfall and information of catchments are required, such as dry weather flow, area and runoff coefficients (see also §3.3.2). The methodology and accompanying software are designed to handle both measurements as simulation results of more detailed models, or a combination of these two. A limited set of available data impacts the modelling process, as already discussed in §2.2.2.

Steps 1 to 3 shown in Figure 3.1 treat the actual configuration of the conceptual model. These three steps form an iterative process. After delineating the topology (see §3.3.1), the required relationships (see §3.3.2-3.3.3) can be identified and calibrated. Finally, the configured modules can be assembled and the model accuracy can be evaluated. If the results are unsatisfactorily, the relationships can be recalibrated, or the modeller can opt for a completely different set of equations. Different topologies can also be examined, for instance by aggregating processes on larger or smaller scales. The developed software tool (see §3.4) automates these three steps greatly during iterations.

3.3 Model components

First, the principles of setting up the conceptual model topology and semantics are discussed. Then, the sewer inflow module is described, followed by an overview of the incorporated model structures that emulate the transport of mass. Table 3.1 summarizes the included model structures (and can be compared to Table 2.1, which lists the model structures applicable for modelling rivers).

3.3.1 Model topology and semantics

Since the storage cell concept underlies the presented methodology, the catchment and sewer system are divided into multiple interconnected virtual storage elements, also denoted as subcatchments (SC). Elements can be lumped on different scales: one such SC can comprise one or more parcels and pipes up to entire districts. Relevant information of the catchment, such as population and areas, are also aggregated to the level of the SC if a rainfall runoff module will be configured in the conceptual model. Hence, the same boundaries are used for both the rainfall runoff and sewer flow calculations.

Figure 3.2b schematizes the conceptualization of the district. Three SCs can be recognized that are interconnected via several pathways. Each flow path (also denoted as a *connection*) depicts one specific structure of the detailed sewer network (e.g. a conduit, pump or other hydraulic structure) for which flow data is present. The whole of connections between two specific SCs is called a *flux*. Thus, such a flux is the accumulation of the water exchange between two SCs. Individual connections can also be aggregated as illustrated in Figure 3.2b: three connections between SCs 1 and 2 are combined into one link (see SC1-SC2 a), while two other flow paths are combined in a different link (SC1-SC2 b). Such an assembly of several flow paths (thus part of the flux) is denoted as a *subflux*. Naturally, if all flow paths between two SCs are merged in one link, the subflux and flux are identical. A flow relationship must be defined for each subflux. Since these relationships are data-driven and the parameters of the conduits or hydraulic structures are not used directly, it is also possible to merge multiple connections of different types (e.g. merging a connection that represent a pump and another that symbolizes a conduit). Combining multiple connections in a subflux reduces the number of relationships in the model. This not only facilitates the calibration effort, it also decreases the simulation time and overall complexity of the model.

Table 3.1: Overview of the different modules, and their usage or advantages. The possibility to enable on/off-switches and flow constraints (such as limiting the flow to a maximum value) is also indicated. Symbols: Q = flow in the sewer system; V = volume; PWL = piecewise linear relationship; TF = transfer function; ANN = artificial neural network; WL = water level; GL = gate level of a controllable hydraulic structure.

Output	Type	Model structure	On/off-switch	Flow constraints	Usage or advantages
V	Continuity equation	-	n.a.	n.a.	Closes water balance in each cell
Q	Fixed flow	Constant value	Yes	n.a.	Simple constant or alternating flows, e.g. fixed pumps
	Static/Dynamic model	PWL	No	No	Simple to complex flows. The optional dynamic component is designed to tackle hystereses.
	TF	TF	Yes	No	Natural flow regimes without variable backwater effects
	ANN	ANN	Yes	Yes	Emulation of complex flow dynamics, including backwater effects and pressurized flows
Inflow	TF	TF	No	No	Modelling of rainfall-runoff
WL	*Not modelled in this approach*				
GL	*Not modelled in this approach*				

Figure 3.2 illustrates this process for one district. The area is partitioned into three SCs. The extent of each cell, and hence also the degree of lumping, depends on the intended use, the desired accuracy, data availability and the behaviour of the system itself. This process is not automated fully to ensure the user can still incorporate expert knowledge, but the modeller is assisted by the created software tool (see also §3.4.1).

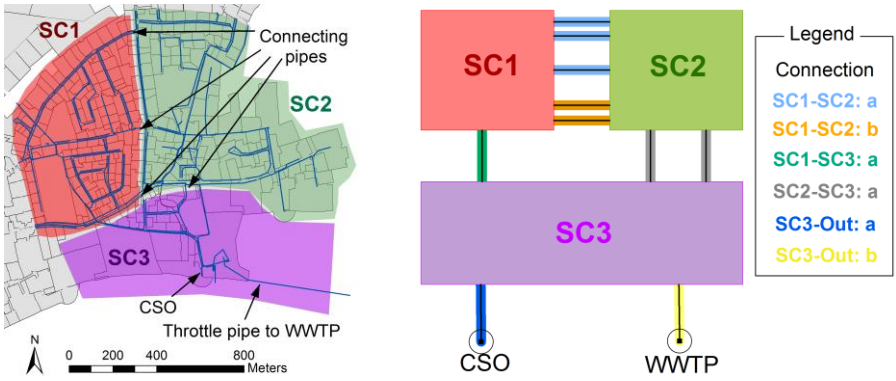


Figure 3.2: a. Detailed sewer network (pipes and hydraulic structures indicated by the blue lines) with an overlay of an exemplary division in three storage cells or subcatchments (SCs; red, green and purple shaded areas); b. Schematization of the conceptual model topology and aggregation of flow paths in several subfluxes.

Note that temporary storage emanating from urban flooding is not considered explicitly in the presented modelling approach. Overland flow between different SCs can likely be represented similarly as subfluxes of sewer pipes, but were not investigated in this research.

3.3.2 Sewer inflow modelling

The conceptual modelling approach incorporates a simple module to calculate the inflow into the sewer system. The approach is highly similar to the methods often incorporated in models and software packages that employ highly detailed spatial catchment information.

The module adopts the same spatial lumping as the conceptual sewer network topology. More specifically, the total contributing catchment is divided into multiple smaller areas or plots (independently from the delineated conceptual model topology). Next, each such plot is designated to one of the defined SCs in

the conceptual model and labelled with a specific land use. A plot can handle rainfall time series, various point inflow hydrographs, a population (and associated dry weather flow series) and several runoff surface types with corresponding areas. Allocating separate rainfall series to every plot also enables the use of data from multiple rain gauges or radar, although the spatial distribution will degrade depending on the detail of the defined topology. Indeed, the plot information and input series are aggregated to the level of the SC. Naturally, this implies that the spatially distributed information of population, input series and surface types within one SC is lost. Each plot in the catchment can incorporate as many surface types as desired, such as pavements, roofs and pervious areas with different characteristics (see Figure 3.3). A separate rainfall-runoff (RR) model is configured for each surface type to transform the rainfall to an inflow hydrograph. During simulations, these calculated RR hydrographs are accumulated with other input sources and dry weather flow to yield the total inflow of a SC.

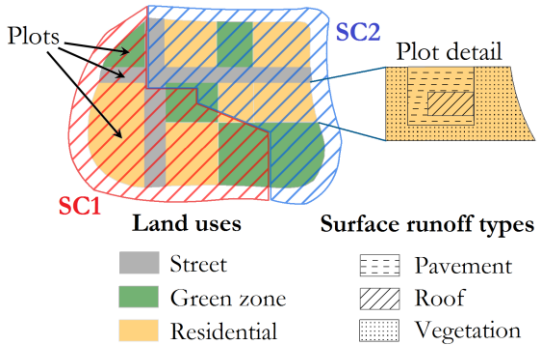


Figure 3.3: Subdivision of an area based on land uses and surface runoff types used by the rainfall-runoff module.

This approach of modelling the inflow can deal easily with changing conceptual model topologies: if a SC is redefined and contains fewer or additional plots, the RR models itself do not have to be recalibrated since the parameters of the RR models are independent of the number of connected plots. It suffices to aggregate the plot characteristics over the newly delimited SCs, which is automatically done by the provided software (§3.4.4).

Many different models exist to describe the transformation from rainfall to runoff. Most of these are variants of the reservoir-routing concept. The Wallingford model, also known as double linear reservoir method and used by default in the InfoWorks CS software, is an example of such approach. It uses

two linear reservoirs in series to represent the storage which is available on the ground and in minor drains, and the delay induced between the peak rainfall and runoff (Innovyze, 2014b). The presented methodology uses a similar approach by using an arrangement of multiple linear reservoirs, each with its own characteristics. Such arrangements of linear reservoirs can be transformed to transfer functions (TF) as proven in the past (e.g. Jakeman et al., 1990; Willems, 2000). Using TFs instead of linear reservoirs has the advantage that many advanced identification and calibration algorithms are readily available. Hence, the RR-models can be configured automatically if runoff results are present. The presented modelling approach and software incorporate the MATLAB System identification Toolbox to set-up the TFs. The steady state gain of the calibrated TF is equal to the runoff coefficient of the surface type and the concentration times of the linear reservoirs can be derived directly from the TF. If runoff results are unavailable, it is also possible to create a TF directly using estimations of the runoff coefficient and concentration time. The RR models do not take losses into account explicitly, but instead apply the rainfall directly to the TF.

3.3.3 Flow calculation

The main part of the developed methodology concentrates on accurate sewer routing. The methodology is designed to predict volumes in SCs and flows between these SCs. The volume is calculated in every SC by closing the water balance explicitly. This is done by combining all subfluxes to and from a SC in the discretized form of the continuity equation. This mechanistic set-up ensures no water is created or lost during simulations. The next sections describe the different incorporated relationships to model subfluxes. The flow through each individual subflux is emulated by one of the presented modules.

3.3.3.1 Option 1: Fixed flow

The first and simplest approach to simulate flows is by combining one or more constant flows and a switch on/off-criterion. This very straightforward concept can be used to mimic the behaviour of pump systems.

Two different binary classification models are included to determine the state of the flux (flow or no flow). The first and most simple approach uses a switch that depends on the volume of one of the SCs present in the model. When the volume of the selected SC surpasses a specified threshold, the specified constant flow is induced until the volume of the SC falls again below the value of the off-

switch. These on- and off-values can be chosen manually, or calibrated automatically based on the presented time series of volumes and flows: the minimal volume of all provided calibration events at the time of flow initiation is used to determine the on-switch, while the minimal volume when the flow stops is used for the off-switch. Secondly, more complicated pattern recognition neural networks can be employed to decide when the specified constant flow should be activated. These neural networks are constructed with sigmoid and softmax transfer functions in the hidden and output layers respectively, and are trained using a scaled conjugate gradient algorithm. The output layer returns for a given input set the probability of equaling or surpassing the selected constant flow. Finally, the modeller chooses from which probability value the constant flow is applied. Another incorporated flow routing method uses neural networks to predict the magnitude of the flow (§3.3.3.4). The reader is referred to that section for further information on the structure, configuration and possible inputs of the neural networks in general. Currently, classification is limited to two categories, but this can easily be expanded to more classes and associated constant flows.

3.3.3.2 Option 2: Piecewise linear static/dynamic (S/D) model

Storage/flow-relationships underlie many hydrologic models. However, most of these relationships in sewer systems are characterized by hysteretic behaviour: during the rising branch of the hydrograph (i.e. with increasing rainfall intensities) a much larger stored volume is present than during the falling branch. Such hysteretic behaviour is typical for systems affected by backwater effects. Vaes (1999) tried to tackle this issue by identifying a static and dynamic storage component. He elaborated this concept for simple sewer systems consisting of one throughflow and one overflow, and incorporated this concept in the REMULI software. Both storage components are linked via piecewise linear relationships to states in the sewer system, such as the throughflow, inflow and overflow. Different sewer systems were modelled using this approach (see e.g. Vaes (1999) and Wolfs et al. (2013)).

The presented modelling approach incorporates a module based on the static/dynamic-theory developed by Vaes (1999), but in a much more adaptable framework. Firstly, due to the flexible model topology, the structure is not confined to handling only one through- and overflow. Secondly, the theory is extended to different sets of inputs. While the original static/dynamic-theory can only associate static and dynamic storage to outgoing and incomings flows respectively, the incorporated approach can link virtually any state (flows,

volumes, water levels or a combination of these, optionally associated with delay elements) to the static and dynamic component via piecewise linear relationships. In this setting, the static component links the flow of the considered subflux to one of these inputs via piecewise linear equations. The optional dynamic component can then be configured to tackle the hysteretic behaviour. Hence, the dynamic segment is calculated during calibration as the difference between the selected input of the static relationship, and the simulated input using the subflux as known time series (i.e. reversed modelling using the earlier defined static component). Next, this calculated dynamic component can again be related to any other desired input set via piecewise linear relationships. Thus, if the volume of the upstream SC and the incoming flow are elected as inputs for the static and dynamic relationships respectively, the proposed approach matches the original concept developed by Vaes (1999).

The piecewise linear relationships are defined manually using interactive figures in the accompanying software tool (§3.4.3). Segmented linear regression with automatic optimization is not incorporated in the current configuration procedure, since the added complexity (such as determining the optimal number of breakpoints) and increased calculation time outweigh its advantages.

3.3.3.3 Option 3: Transfer function (TF)

TFs were already described in §2.3.1.4 in the discrete domain. This section approaches the TF theory from a continuous time perspective and shows the link between linear reservoir theory and TFs. Using the continuous TF formulation for urban drainage modelling and the discrete for river modelling has grown historically in the research group. Naturally, the developed conceptual modelling approaches are compatible with both formulations and can be interchanged freely.

Linear reservoir models are a popular technique to route flow in sewer systems. Such approaches reason from the continuity equation in combination with the assumption of a simple linear relationship between the outgoing flow of the (virtual) reservoir and its storage $Q = V/k$ with k a time constant. Elaboration yields the following differential equation:

$$k \cdot sQ = I - Q \tag{3.1}$$

with *s* the derivative operator $s^r = d^r/dt^r$ and *I* the incoming flow upstream. Rewriting this equation according to the ordinary notation of TFs and multiplication with a steady state gain *G* yields:

$$\frac{Q}{I} = \frac{G}{1 + k \cdot s} \tag{3.2}$$

Analyzing TFs in the continuous domain allows for a much more transparent interpretation in physical terms than (equivalent) discrete-time models. The inverse Laplace transformation of this TF yields an analytical expression of the impulse response (IR), which gives after convolution with the incoming flow *I* the output *Q*. The IR defines how the system reacts to an instantaneous pulse input at a certain time τ which is zero elsewhere. This analytical IR is characterized by an exponential decay that is directly related to the earlier defined time constant *k*: at time $t = \tau + k$, the response recedes to $\exp(-1)$ or circa 37% of the initial response at time τ. Alternatively, time constant *k* reflects the shift of the centroid (i.e. the geometric center) of the incoming and outgoing flows. It is interesting to note that this IR can also be interpreted as the in hydrology widely known Instantaneous Unit Hydrograph (IUH; see also Yang and Han, 2006). Thus, there is a direct link between simple first order TF systems (Eq. 3.2) and the IUH.

Instead of manually testing different arrangements of linear reservoirs with varying parameters, numerous algorithms are readily available to identify the most suitable model structure of the TF and calibrate its parameters. As explained below, a direct link between such TFs and the linear reservoir theory remains in most cases. Equation 3.3 gives the general expression for TFs of higher orders.

$$\frac{Q}{I} = \frac{b_m s^m + b_{m-1} s^{m-1} + \cdots + b_1 s + b_0}{a_n s^n + a_{n-1} s^{n-1} + \cdots + a_1 s + a_0} \cdot exp\,(-T_d s) \tag{3.3}$$

with *a* and *b* real coefficients and T_d the delay of the TF. In this research, the order of the TF is limited to maximal three (i.e. 3 zeros in the denominator, also denoted as "poles") and one zero in the numerator ("zero" of the TF). This range of structures should be versatile enough to emulate most systems in hydrology. Higher orders lead to more complex systems, which are also more difficult to interpret in physically meaningful terms. Note that the equivalent TF formulation in discrete form is given by Eq. 2.1 (which additionally contains an error model).

Since a TF represents a differential equation, inspecting the position of its zeros and poles in the (imaginary) space gives insight into the system response. In particular, the system poles define directly the homogeneous response of the system as a summation of exponential functions of real or complex values. Since the poles are such an important characteristic of the TF, two different approaches are incorporated in the methodology to configure TFs with special attention to the pole estimation. Firstly, the roots of the TF can be constrained to be real. These TFs can be easily decomposed into smaller (first order) TFs using partial fraction expansion (see Kuo and Golnaraghi (2002) and Young (2011) for a more comprehensive text), such as for example the TF of Eq. 3.2. Thus, these higher order TFs can be broken down to smaller and easier interpretable systems. This TF can also represent the superposition of several linear reservoirs in series, parallel or a combination of both. Multiplying the TF of individual components represents a serial connection, while the addition of TFs describes a parallel configuration. Thus, although a TF appears to be "black box", it can actually be employed to gain insight into the system's dynamics. They have a credible and adaptable analogy with linear reservoirs (e.g. Jakeman et al., 1990). This approach also fits into the data-based mechanistic (DBM) modelling philosophy where models are identified and calibrated based on data, but are only accepted to be useful if they have a legitimate (but sometimes unconventional) physical interpretation (e.g. Young, 1993; Young and Beven, 1994; Young et al., 1997; Lees, 2000). In fact, many of the popular conceptual models based on linear reservoir theory such as Nash (1957), Diskin (1964) and Kulandaiswamy (1964) are special cases of a TF as proved earlier (e.g. Spolia and Chander, 1974). Wolfs et al. (2013) showed that the original static/dynamic REMULI model (see §3.3.3.2) involving volume can also be translated to a TF when combined with the continuity equation. On the other hand, the optimization can allow complex poles. Since all coefficients in the denominator are real, such poles will come in conjugate pairs. These systems are also referred to as "underdamped" systems and are often written in the following form (in case of a 2nd order TF with no zero):

$$\frac{Q}{I} = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n + \omega_n^2} \cdot exp\,(-T_d s) \tag{3.4}$$

with ω_n the natural frequency of oscillation of the TF and ζ the damping factor. Note that for $\zeta \geq 1$, the TF will be characterized by real roots, and decomposition into first order systems is possible as explained above. TFs with $\zeta < 1$ have

complex (conjugate) roots and the position of these roots in the complex plane determines the homogeneous response. If these roots have negative real parts, the TF will react stable, otherwise the system's response will show oscillations that diverge. Purely imaginary poles result in sustained oscillations. The sinusoidal response of TFs with complex poles, either decaying or diverging, is unlikely to occur in real hydrological systems and thus might indicate an overparameterized system. Interpreting the system as an arrangement of linear reservoirs is not possible.

The steady state gain *G* of the TF can be optimized or constrained to unity. A steady state gain of unity implies conservation of mass if the input of the TF is a flow. The modelling approach and software tool allows using up to three TFs in parallel with different input sets to calculate a subflux. As in the S/D-model (see §3.3.3.2), such input set can be very flexible and consist of (a combination of) water levels, flows and volumes present in the sewer system. Identification of the most suitable TF structure (up to three poles and one zero) is done based on the Akaike Information Criterion (Akaike, 1974). This criterion provides a general measure of model quality by analyzing the goodness of fit versus the model complexity (i.e. the number of parameters in the TF). Finally, the TF can be combined with the switch on/off-criterion described in §3.3.3.1. This enables the TF to imitate discontinuities in predictions (such as when the overflow starts or stops spilling abruptly).

3.3.3.4 Option 4: Artificial neural network

As discussed in §2.3.1.2, simplified relations with a fixed or rather rigid model structure often have difficulties in emulating more complex flow dynamics, such as variable backwater effects, pressurized or reverse flows. It is impossible to enumerate even the majority of the influential factors of the flow rate. Machine learning techniques, however, are very flexible and can actually learn from the data provided during configuration. These structures can adapt themselves and can therefore often emulate more complex patterns and relationships.

Therefore, ANNs are also incorporated in the modelling approach for urban drainage systems to estimate the flow through subfluxes. More specifically, feedforward networks with one hidden layer and a selectable number of hidden nodes can be incorporated with up to three inputs sets. These input sets can be defined similarly as for the S/D-model (see §3.3.3.2) and TFs (see §3.3.3.3). One of the main problems encountered when training ANNs is the risk of overfitting

the structure. In such cases, the model succeeds in mimicking the calibration data accurately, but fails to generalize adequately and simulating new data yields erroneous predictions. Two techniques are incorporated in the methodology and accompanying software tool to improve the generalization capability of ANNs. As for the ANNs in the developed conceptual modelling framework for river systems (§2.3.1.2), networks are by default configured using the early stopping technique in the CMD software. This implies that the data of selected calibration events is divided randomly into two subsets: a training (80% of the calibration data) and a validation set (the remaining 20%). The training set is used to optimize the network weights and biases, while the error on the validation set is monitored during this training process. If the validation error increases for a specified number of iterations, training is stopped. Finally, the performance of the ANN can be checked using the independent validation events (often denoted as the "test set"). Secondly, the modeller can use ensembles of configured ANNs during simulations. This technique averages the output of a selected number of ANNs and is particularly effective when the data set is small or noisy. Each ANN in such ensemble is configured using newly created training and validation sets and different initialized weights and biases.

ANNs are continuous mathematical models, meaning that they have difficulties with modelling discontinuous system behaviour, such as sudden jumps in flow predictions. Therefore, ANNs can also employ the switch on/off modules described in §3.3.3.1 to determine when flow initiates and ceases. Similarly, an additional switch module can be incorporated to limit the flow to a specified maximum. In such case, the switch determines whether this maximal value should be used as outcome of the flow prediction model, or the simulated value of the ANN. Finally, the outcome of the flow module can be constrained to yield solely positive or negative values. This mechanism was implemented to emulate flap valves. If the ANN predicts a flow in the inhibited direction during simulations, the value is reset to zero. Employing these switches and constraints permits to focus on non-zero and non-maximum flows entirely during training, resulting in more accurate ANNs with fewer hidden nodes.

3.4 Software

As for the approach to conceptually model river systems, the methodology presented in this chapter is also accompanied by a semi-automatic software tool. The software speeds up configuration and guides the modeller through the

different model options. The software is also named CMD, but is coded separately in MATLAB from the CMD tool for rivers (see §2.5). This facilitates maintenance, development and future interfacing or integration with other tools. The CMD tool for urban drainage modelling is equipped with an extensive GUI. Again, the conceptual model itself is constructed in C programming language due to its computational efficiency and straightforward interfacing options.

The software is constructed around a framework that manages and organizes all data in the background. Different GUIs can be invoked from this framework, each representing one of the main functions of the software tool. The outcome of the GUIs is mutually passed without intervention of the user. Its main functions match the different steps of the model configuration process:

1. The division of the (detailed) sewer network in multiple interconnected SCs (see also §3.3.1)
2. Aggregating connections between these different SCs into so-called subfluxes (§3.3.1)
3. Handling calibration and validation events and data
4. Identification and calibration of model structures for each defined subflux (§3.3.3)
5. Configuration of model structures and parameters that determine the sewer inflow (§3.3.2)
6. Model assembly and simulation

The software's features for each of these functions are briefly described hereafter. The first two functions determine the conceptual model topology. For practical reasons, these functions are accessed via separate GUIs in the software, but are addressed simultaneously in the following text.

3.4.1 Function 1: Defining the model topology

The first main function is designed to define the model topology (see also §3.3.1 for discussion on the model conceptualization and semantics). A sewer network of a detailed InfoWorks CS model (or a model configured in different software as long as the network file has a similar template) can be imported. Next, a plot is shown with all links and the different hydraulic structures and outfalls can be identified. Different types of sewer elements can be distinguished, such as storm, combined and other systems. This representation of the detailed system allows the modeller to allocate nodes to different SCs easily: nodes can be selected by

simply drawing polygons using the cursor, or by selecting a node and calculating up- and downstream traces. The software identifies such traces based on the up- and downstream node definitions of each link. These traces can also be used to designate non-allocated nodes automatically to defined SCs. Finally, water level boundary nodes can be selected that will be used to estimate the flows during simulations. This functionality is crucial to capture the interaction with external systems, such as rivers and ponds. Note that the presented approach does not aim to simulate these water levels (see also §3.2.1). Figure 3.4 shows the GUI that controls the described function for allotting nodes to specific SCs.

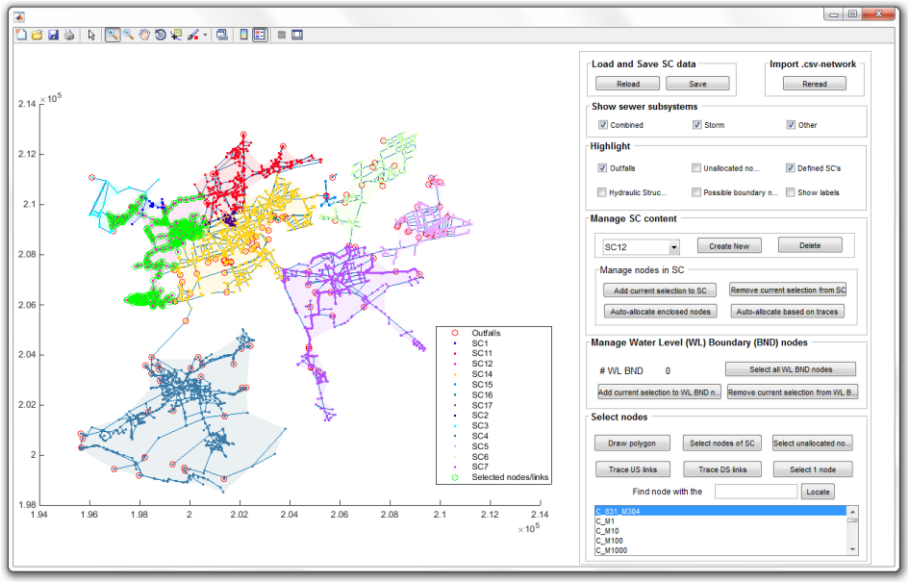


Figure 3.4: GUI for dividing the detailed sewer network in subcatchments (SCs).

Once the detailed network is imported and the different SCs of the conceptual model are delineated, the software interprets which SCs are interconnected and via what flow paths. Next, the user can merge different flow paths into one subflux. Later, a model structure will be identified and calibrated for each subflux (see Function 4, §3.4.4). Consequently, merging links leads to a reduced number of model structures and parameters, which in turn results in a lower calibration effort, a more robust model and a shorter calculation time. Naturally, different hydraulic structures, such as weirs, sluices and regular conduits, can also be combined in a single subflux since the parameters of these hydraulic structures are not directly used in the emulation model structure. As with the delineation of

SCs, the level of aggregation depends on the model behaviour and a trade-off between the desired accuracy and calculation speed of the final conceptual model.

3.4.2 Function 2: Manage event data

After the model topology is specified, data can be imported from databases to configure the model structures that emulate the flow through every subflux. These databases should be constructed in a spreadsheet format (such as *.xls*, *.xlsx*, *.csv* etcetera), contain the names of the variables and associated time series, and the time interval. Thus, these databases can be composed easily or even exported directly from the InfoWorks software. The flow through the links of all subfluxes and inflow of every node in the sewer network are imported. The water levels of selected boundary nodes are also loaded. Next, the tool calculates the volume of each SC. In case the model topology or database content is altered by the modeller and subflux models are already configured, the software automatically removes affected calibrated model structures and warns the user. This ensures no model anomalies can originate. The associated GUI gives a comprehensive overview of the loading process and allows adding, removing or reloading different events.

3.4.3 Function 3: Calibrate subflux flow models

Once the model topology is set up and the required data are imported, a model can be calibrated for each specified subflux. The user can choose from four incorporated model structures (fixed flow (§3.3.3.1), piecewise linear S/D model (§3.3.3.2), transfer functions (§3.3.3.3) and neural networks (§3.3.3.4)), together with additional integrated modules such as binary classification models. The lay-out and functionality of the GUI change depending on the elected model structure, which ensures that the modeller is led efficiently and in a step-wise manner through the configuration of the chosen model structure. A separate pane was designed to compose the desired input set for the model structure. All necessary algorithms to identify and calibrate the most appropriate relationships of the elected module and associated parameters are incorporated in the software itself, and can be controlled via the GUI. Different events can be selected for calibration and validation. After configuration, the emulated flow is shown together with the original calibration and validation time series, together with calculated goodness-of-fit criteria. Finally, configured neural networks are converted into one, two or three dimensional lookup tables (depending on the number of inputs). Such low dimensional tables can be visualized easily, thereby

allowing the user to assess the extrapolation behaviour and the physical soundness of the neural network. In addition, this transformation facilitates the efficient incorporation in the final calculation scheme written in C programming language.

3.4.4 Function 4: Sewer inflow modelling

The GUI to model the sewer inflow can function isolated from the other GUIs. The user can import rainfall time series and additional incoming flows, together with network characteristics such as the population, defined land uses, surface runoff types and associated areas per parcel. The required formats for these data are similar as described earlier (§3.4.2) and these files can again be directly exported from the InfoWorks software.

A separate RR model has to be configured for each defined surface runoff type (see also §3.3.2). The tool contains a database that organizes all previously saved RR models. The modeller can pick one of these configured structures or choose to calibrate a new RR model. For the latter, target values have to be imported as well since the calibration procedure is data-driven. Finally, the user has to specify the dry weather flow per capita.

Once the user has defined the conceptual model topology, the tool can be applied to assemble all RR model structures and corresponding parameters to simulate the inflow of each defined SC. The conceptual topology is passed internally to map the detailed parcel information onto the defined SCs. The different rainfall series and characteristics of the plots (such as the area per runoff surface type) are lumped over the entire SC and finally, one inflow series is calculated by the software. Note that the implemented approach does not require redefining the sewer inflow model if the conceptual topology is altered. Instead, the tool can simply aggregate all parcel information according to the newly defined SCs.

3.4.5 Function 5: Model assembly and running simulations

Finally, all configured modules can be assembled into one script written in C programming language. A GUI provides a clear and comprehensive overview of the subfluxes that have already been calibrated, together with additional details such as the elected model structure and associated modules, used input set, goodness of fit and the time when the structure was last modified. It is also

possible to visually inspect the input(s) and subflux model outcome, and compare it with the original time series.

The user can select the precise content of the conceptual model through the provided GUI. One can exclude several SCs from the final calculation scheme. Relevant flows (such as subfluxes from SCs that are excluded from the model to neighbouring SCs that are incorporated in the calculation, or subfluxes that are not represented by a model yet) are passed as boundary information to the model. This ensures that the model can still be simulated correctly and allows the user to inspect error propagation through the model rigorously. Finally, the user can specify a desired calculation time step. Note that the TFs are estimated in the continuous domain, and its parameters depend on the sampling interval of the calibration data (§3.3.3.3). Therefore, these TFs are converted to the discrete domain using a zero- or first-order hold method taking the specified time step of the final calculation scheme into account.

Once the model is written, the software compiles the C-script and simulates the included events. It is possible to create an automatically generated report which lists all variables that are being calculated in the model, together with all boundary data and a detailed overview of the model topology. For each calculated variable and every simulated event, the Nash-Sutcliffe efficiency is calculated by comparing the simulated and original time series.

3.5 Case study

3.5.1 Study area and events

The set-up of the conceptual model of the sewer system of the villages Oostakker and Sint-Amandsberg, two districts of the city of Ghent in Belgium, is described to illustrate the presented modelling approach and software tool. The area covers about 2.750 hectares and is highly urbanized (38% impervious area), with a population equivalent of 43.626. Around 84% of the system is configured as a combined system, 15% storm type and the remaining is of the conveyed sanitary type. The area is very flat, with an average slope of 3.5‰. This results in a very looped network as illustrated by Figure 3.5a. A detailed InfoWorks CS model was available of the examined sewer system. This model was set-up earlier and combines detailed information of the catchment area, implemented pump settings and controls, and location, geometry and material of the pipes. The

network counts in total 6.025 conduits, 182 hydraulic structures (such as pumps, weirs, sluices and orifices) and 5.855 manholes. The system releases water to receiving water bodies via 39 outfall nodes. Calibration was performed according to the guidelines delineated in the Hydronaut procedure, which is a methodology that is commonly used to configure hydrodynamic sewer models (Aquafin, 2005).

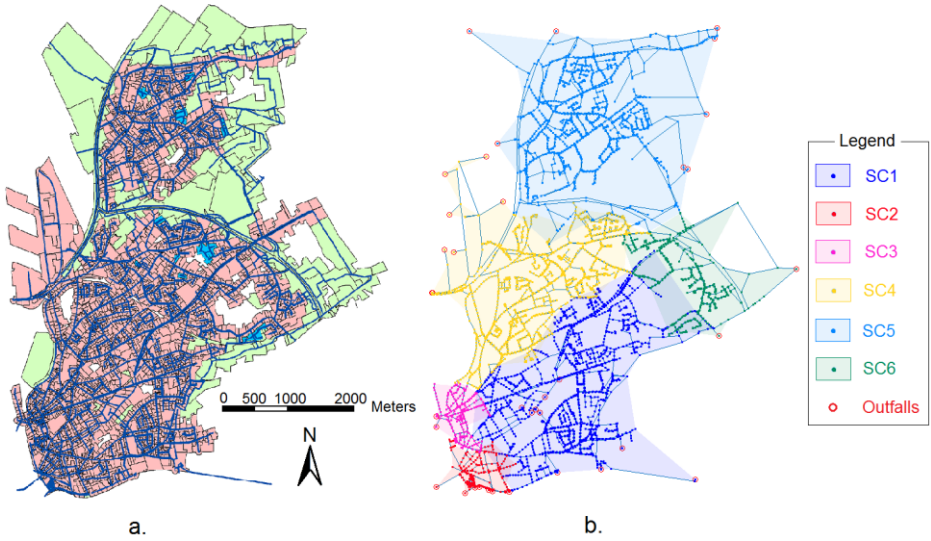


Figure 3.5: Detailed InfoWorks CS model of the case study area (a) and the conceptual model topology (b).

3.5.2 Conceptual model configuration

The aim of the conceptual model is to simulate the outgoing discharges (CSOs, pumps, etc.) accurately for given rainfall time series with an absolute minimal calculation time. Thus, the conceptual model includes a rainfall runoff and sewer transport component. Different synthetic storms are simulated in the detailed InfoWorks CS model with frequencies of occurrence of 10 (denoted as 'f10') and 7 ('f07') times per year, and return periods of 2 ('t02'), 5 ('t05'), 10 ('t10') and 20 ('t20') years. The simulation results of every element in the network are stored every 300 seconds and used to configure the conceptual model. Setting up a conceptual model is an iterative procedure (see also §3.2.2), but only the final retained configuration is discussed below.

3.5.2.1 Conceptual model topology

The first step in the set-up of the conceptual model is the definition of the topology (§3.3.1). Since the objective of the model is a minimal calculation time, processes are lumped extensively. Therefore, the entire system was divided in 6 SCs, as shown in Figure 3.5b. This division was obtained by calculating the upstream flow paths of each outfall node using the built-in functions of the software, and clustering the nodes of similar paths in a SC. Due to the looped character of the detailed network, it was impossible to isolate SCs completely. Instead, the identified SCs are linked via multiple flow paths, but the followed approach based on upstream trace calculation minimizes the interconnectivity. Finally, any remaining nodes were allocated to SCs by investigating the system dynamics. Experiments indicated that fewer SCs would aggregate processes on too large scales, leading to inaccurate predictions.

Once the SCs are delineated, the flow paths between the SCs and to outfall nodes can be merged to minimize the number of calculation nodes, and thus reduce the calibration effort and simulation time. Figure 3.6 shows the 9 different flow paths from SC2 to the sewer outfall nodes for the 'f07' storm, each represented by a line. One of these links is located just downstream of a pumping station that spills into the Scheldt River (Figure 3.6b), while the other links represent conduits or flap valves that release water gravitationally to the receiving river and have similar flows (Figure 3.6a). Naturally, the flow through the pumping station differs significantly from the other flow paths, and therefore, this link was isolated from the others and modelled in a separate subflux. Because all other flow paths behave similarly to each other and it is not necessary to know the precise spatial distribution of the sewer outfall discharges, these 8 paths were merged into a single subflux. A similar procedure was followed for all other flow paths to outfalls.

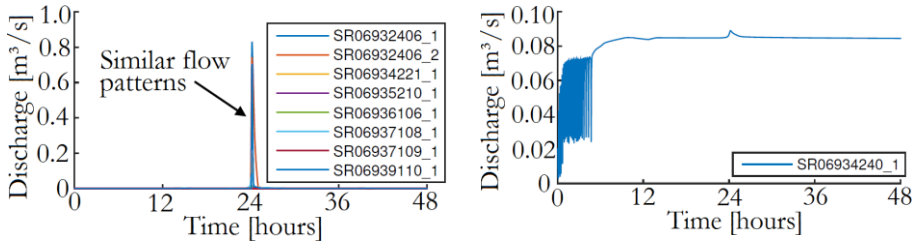


Figure 3.6: Aggregation of the flow paths from SC2 to the sewer outfall nodes in two subfluxes. Different flow paths with similar discharge series are aggregated in

a single subflux (a), while another flow path with distinctive response characteristics is modelled in a separate subflux (b).

Because of the thoughtful demarcation of SCs, the interlinking between the SCs in the conceptual model remains limited. Table 3.2 shows the number of flow paths between each SC. All flow paths between two SCs were merged into one subflux, except the connection between SC1 and SC6, where the flow connections were modelled by two subfluxes.

Table 3.2: Number of flow paths between the different SCs and sewer outfalls, together with the corresponding number of defined subfluxes between brackets.

	SC1	SC2	SC3	SC4	SC5	SC6	Outfalls
SC1	-	3 (1)	4(1)	9 (1)	0	6 (2)	10 (2)
SC2	3 (1)	-	1 (1)	0	0	0	9 (2)
SC3	4 (1)	1 (1)	-	2 (1)	0	0	2 (2)
SC4	9 (1)	0	2 (1)	-	8 (1)	2 (1)	9 (1)
SC5	0	0	0	8 (1)	-	1 (1)	9 (1)
SC6	6 (2)	0	0	2 (1)	1 (1)	-	3 (1)
Outfalls	10 (2)	9 (2)	2 (2)	9 (1)	9 (1)	3 (1)	-

Finally, the flow through one of the outfalls of SC6 was significantly influenced by the water level of the receiving river. Therefore, this water level was incorporated as boundary condition in the conceptual model. The sewer model can later on be interfaced with a river model that estimates this water level, and thereby model the sewer-river interactions.

3.5.2.2 Model structures and parameters

After definition of the conceptual model topology, the identified subfluxes can be calibrated. Each subflux is modelled individually. The four incorporated model structures are described in §3.3.3. The following text briefly describes the configuration of four selected subfluxes to illustrate the functioning of each model structure. Other subfluxes were configured similarly. The events 'f07', 't02' and 't20' were used for calibration, and 'f10', 't05' and 't10' for validation.

A piecewise linear S/D model (§3.3.3.2) was set-up to emulate the earlier identified subflux from SC2 to one of the outfalls that represents a pumping station (see §3.5.2.1 and Figure 3.6b). The detailed InfoWorks CS model

incorporates a vertical sluice, orifice and pump in series to mimic the settings of this pumping station. The volume of SC2 was elected as input for the static model. It is clear that the pumping rate is closely related to this storage, and even no hysteretic behaviour can be noticed (see Figure 3.7a). Therefore, it was not necessary to incorporate the optional dynamic component. Note that the identified static relationship is extrapolated linearly beyond the calibration range. Thus, the modeller can also estimate the extrapolation behaviour of the relationship during calibration. The simulation results of this model structure in offline mode (i.e. the inputs are copied from the calibration data set and not yet simulated by the conceptual model itself) are shown in Figure 3.7b for the event 'f07'. Naturally, the rapid oscillations are not captured by the simplified static model because it ignores the non-univocal behaviour around the very steep upward part of the static model, but this is also not the aim of the model. Instead, the conceptual model follows an averaged trend through the rapidly fluctuating flow of the detailed InfoWorks CS model and reacts more stable. The higher flows are also captured accurately.

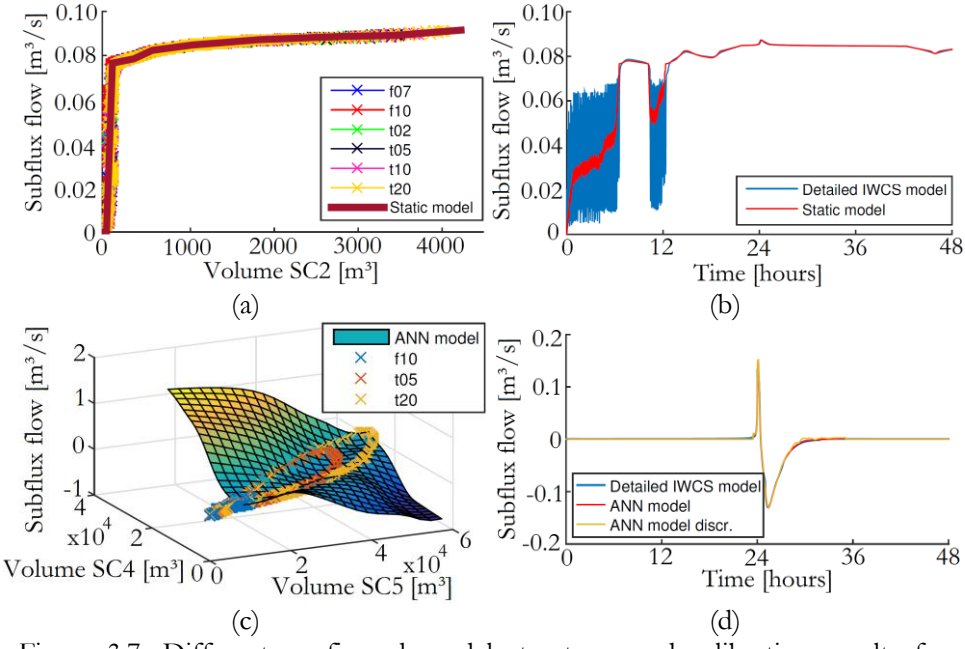


Figure 3.7: Different configured model structures and calibration results for selected subfluxes. (a) Static model of subflux "SC2-b". (b) Calibration result of the static model of subflux "SC2-b" for event 'f07'. (c) Neural network (ANN) of subflux "SC4-SC5-a". (d) Calibration result of the ANN of subflux "SC4-SC5-a" for event 't05'.

Next, an ANN (§3.3.3.4) was configured that mimics the subflux between SC4 to SC5, which is the aggregation of 8 different flow paths. The flow is heavily influenced by backwater effects and the flow direction changes abruptly during simulations with the detailed InfoWorks CS model. ANNs are ideally suited to capture such complex flow dynamics. The volumes of SC4 and SC5 were chosen as inputs, since these will arguably form the main drivers of the flow. First, a simple switch on/off criterion was calibrated based on the volume of SC4 that determines whether flow should be initiated or not (§3.3.3.1). Next, a neural network with 10 hidden nodes was trained using early stopping and transformed to the grid shown in Figure 3.7c. This conversion allows the user to visually inspect the outcome of the model for values outside the calibration range, and thus assess the extrapolation capability. During simulations with the conceptual model, input values outside of the lookup table are mapped to the closest point on the grid to control extrapolation. Figure 3.7d shows the almost perfect match with the calibration results for the 't05' event.

A subflux from SC3 to one of sewer outfalls is modelled in the detailed InfoWorks CS model by a simple pump with a constant flow rate. Naturally, a fixed flow module (§3.3.3.1) was configured to imitate the flow through this pump. A switch on/off criterion that depends on the stored mass of water in SC3 managed to reproduce the operation of the pump accurately, and thus it was superfluous to incorporate the more advanced binary classification neural network module. The calibration result and original series for event 't05' are presented in Figure 3.8a.

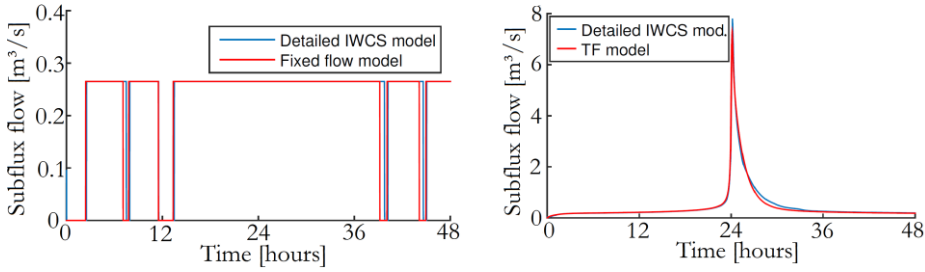


Figure 3.8: Different configured model structures and calibration results for selected subfluxes. (a) Calibration result of the fixed flow module of subflux "SC3-a" for event 't05'. (b) Calibration result of the transfer function (TF) module of subflux "SC4-a" for event 't05'.

A TF model (§3.3.3.3) was identified and calibrated to emulate a subflux from SC4 to sewer outfalls. All flows from neighbouring SCs (defined positive towards SC4) were summed together with the incoming flow due to runoff and other inflow hydrographs to form the input of the TF. Different TF orders were calibrated, and finally a TF with a single pole and zero was retained. This TF was selected by comparing its Akaike Information Criterion and Nash Sutcliffe efficiencies with alternative TFs. The TF is characterized by a recession constant of 4196 s and has no additional delay, which corresponds with the centre of gravity shift of the incoming and outgoing flow hydrographs. Again, the good agreement between the TF module and detailed InfoWorks CS simulations results can be noticed in Figure 3.8b.

After the subflux models were configured, the Qnode inflow model was set-up. The Qnode inflow model combines the rainfall runoff (RR) of the catchment with dry weather flow (DWF) and other selected inflow hydrographs. A constant DWF, equal to the average of the flow in the InfoWorks CS model, is applied in the conceptual model. Incorporating the precise time variation of the DWF is unnecessary, since this only has a negligible effect on the magnitude of the overflow fluxes. The conceptual model mimics this RR module of the available detailed InfoWorks CS model. This model contains 4 different surface runoff types (see also §3.3.2), and consequently, a separate TF was identified and calibrated based on simulation results of the detailed models for each such runoff type. Finally, the different events were simulated and the inflow for each defined SC was calculated. The Nash Sutcliffe efficiencies of those simulations all surpass 99% when compared to the simulation results of the detailed models.

3.5.3 Model accuracy

The configured conceptual model was then used to simulate all six investigated events. It is impossible to evaluate the model's performance of all calculation nodes due to the extensiveness of the model. Instead, the simulated volumes of the identified SCs and the overflow fluxes are compared to the simulation results of the detailed InfoWorks CS model. The Nash-Sutcliffe efficiency (NSE, Nash and Sutcliffe, 1970) is calculated for every selected flow and volume (see Eq. 2.7). An NSE-value of unity indicates a perfect fit. Table 3.3 shows the obtained NSE-values for all SCs. All events with larger return periods are emulated accurately, with NSE-values ranging from 0.82 to 1, and an average value of 0.95 for events with a return period of 2 years and more. For events with higher frequency of occurrence, the conceptual model results deviate somewhat from the results of

the detailed model, with NSE-values of a single SC dropping to around 0.6. This can be expected however, since the optimization techniques focus more on the events with larger return periods during model set-up. In addition, SC3 is characterized by smaller volumes, and is consequently also more sensitive to deviations. Even then, the peak volumes are estimated accurately (Figure 3.9a). These peak volumes are the main drivers for the sewer overflow fluxes. Since this case study primarily aims at predicting these overflow fluxes accurately, it is especially important to capture the peak volumes in the SCs correctly. To improve the prediction of storms with higher frequency, additional calibration events can be incorporated with lower flows, or an alternative objective function could be selected that penalizes prediction errors on high flows less.

Table 3.3: Nash-Sutcliffe efficiencies expressing the fit of the simulated volumes in the conceptual model to the results of the detailed InfoWorks CS model for all defined SCs and events ('C' and 'V' indicate calibration and validation events respectively).

	f10 (V)	**f07 (C)**	**t02 (C)**	**t05 (V)**	**t10 (V)**	**t20 (C)**
SC1	0,97	0,98	0,90	0,93	0,94	0,95
SC2	0,96	0,96	0,94	0,93	0,94	0,93
SC3	0,61	0,67	0,90	0,92	0,92	0,92
SC4	0,75	0,81	0,98	0,97	0,97	0,97
SC5	0,97	0,98	1,00	1,00	1,00	1,00
SC6	0,99	1,00	0,98	1,00	1,00	1,00

Table 3.4 shows the NSE-values of all subfluxes leading to sewer outfalls. In addition, the relative volume error (Δ) is listed of the cumulated flow of the simulation results of the conceptual model compared to the results of the detailed InfoWorks CS model, together with the simulated spilled volume of the InfoWorks CS results (Δ').

Table 3.4: NSE-values indicating the fit of all subfluxes in the conceptual model leading to sewer outfalls to the results of the InfoWorks CS model for all events ('C' and 'V' indicate calibration and validation events respectively). The relative difference (Δ) of the cumulated flows of the conceptual model and the detailed model are shown, together with the spilled volume in the detailed model (Δ').

		f10 (V)	**f07 (C)**	**t02 (C)**	**t05 (V)**	**t10 (V)**	**t20 (C)**
	NSE	0,97	0,98	0,99	0,99	0,99	0,99
SC1 – a	Δ	-3%	4%	3%	3%	3%	3%
	Δ'	18.423	19.368	59.468	73.913	85.453	97.045
	NSE	0,95	0,95	0,94	0,94	0,93	0,90
SC1 – b	Δ	0%	-1%	-2%	-2%	-2%	-2%
	Δ'	148.823	156.523	217.381	223.315	227.529	231.032
	NSE	0,79	0,85	0,98	0,97	0,98	0,98
SC2 – a	Δ	-61%	-49%	-11%	-9%	-8%	-6%
	Δ'	419	645	3.677	4.897	5.820	6.781
	NSE	0,80	0,79	0,79	0,79	0,79	0,78
SC2 – b	Δ	-1%	-1%	-1%	-1%	-1%	-1%
	Δ'	12.371	12.925	14.077	14.112	14.139	14.160
	NSE	0,55	0,50	0,24	0,34	0,34	0,34
SC3 – a	Δ	0%	1%	1%	0%	0%	0%
	Δ'	31.113	32.068	39.729	40.317	40.300	40.365
	NSE	1,00	1,00	0,98	0,99	0,99	0,98
SC3 – b	Δ	0%	0%	1%	3%	2%	1%
	Δ'	0	0	5.313	6.792	8.248	9.718
	NSE	0,96	0,97	0,99	1,00	1,00	0,99
SC4 – a	Δ	2%	2%	-1%	-2%	-2%	-1%
	Δ'	31.552	35.564	71.043	79.708	86.189	92.558
	NSE	0,99	0,99	1,00	1,00	1,00	1,00
SC5 – a	Δ	0%	0%	0%	0%	0%	0%
	Δ'	47.719	55.458	112.107	126.340	137.117	148.068
	NSE	0,84	0,84	0,87	0,90	0,91	0,94
SC6 – a	Δ	2%	1%	-6%	-4%	-3%	-1%
	Δ'	-25.890	-24.714	-14.991	-11.861	-9.479	-7.072

Most NSE-values surpass 0.9, indicating a good fit. The predictions of subflux "SC3-a" have clearly inferior NSE-values compared to the other flows. This subflux represents a pump with a fixed discharge. Although the conceptual model manages to simulate the cumulated flow very accurately (with a relative error of maximum 1%, see Table 3.4) and imitate the on and off switching of the pumps reasonably well (see Figure 3.9c), the NSE-values vary between merely 0.24 and 0.55. The flows of subfluxes "SC2-b" (see also Figure 3.7a and b) and "SC6-a" show rapid fluctuations in the simulation results of the detailed InfoWorks CS model. The conceptual model does not emulate these fast oscillations, but yields more damped and averaged flows (see also Figure 3.9d). Again, this results in lower NSE-values, while the cumulated flows show only minor deviations from the simulation results of the detailed model.

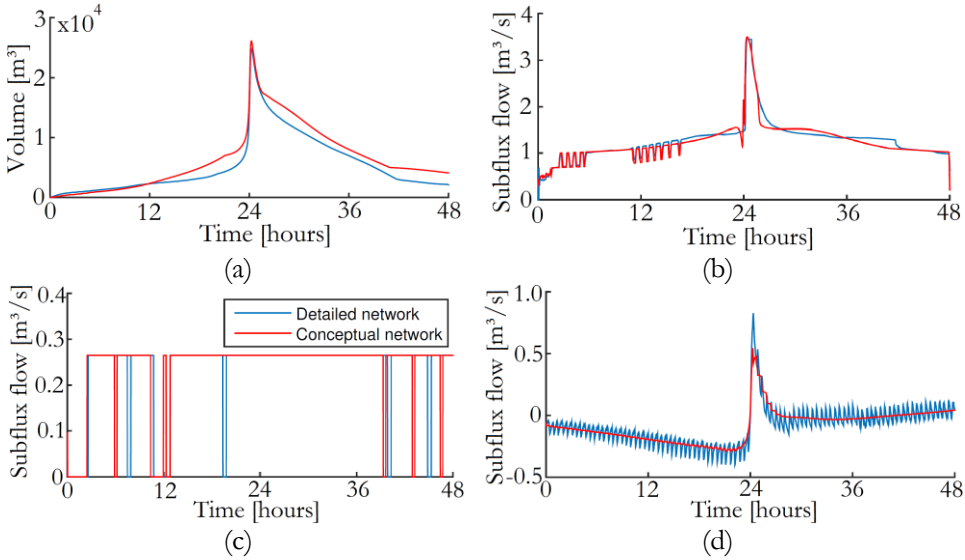


Figure 3.9: Simulated volume of SC1 (a) and flows in the detailed and conceptual model for subfluxes "SC1-b" (b), "SC3-a" (c) and "SC6-a" (d) for the event 't02'.

The conceptual model has a calculation time of merely $2 \cdot 10^{-3}$ s for a two-day event with a time step of 60 s using a single core of an i7 processor clocked at 3.40 GHz. This is almost 10^6 times faster than the detailed InfoWorks CS model, which takes approximately 30 minutes to simulate the same event and uses multiple cores. Due to this vast speed gain, the conceptual model can be employed in numerous applications requiring a short calculation time.

3.6 Conclusions

This chapter presents a novel conceptual modelling methodology for sewer hydraulic computations. It was designed to address the requirements for present and future urban water management outlined in §1.2.1, and reasons from the same principles and modular approach as the developed methodology to conceptually model rivers (Chapter 2). The main differences with the approach for rivers presented in the previous chapter are (1) a modified conceptual model topology, (2) only flows and volumes are modelled, hence water levels and gate levels are not simulated, (3) different model structures are employed to calculate flows and (4) the inflow into the sewer system is also incorporated in the methodology. These changes ensure that the modelling approach is tailored to modelling urban drainage systems, and can deal with the complex and fast dynamics often encountered in sewers, such as pressurized and reverse flows, and variable backwater effects. The novel methodology is highly flexible and computationally very efficient. Processes can be lumped on different scales, making it possible to focus only on the dominating ones that are relevant for the intended application. The modular design allows picking and combining various incorporated model structures in a mechanistic setting. The different model elements are combined in an adaptable calculation scheme programmed in C language, and can therefore be integrated easily with other modules.

An extensive and semi-automatic software tool (Conceptual Model Developer, CMD) was developed that guides the user through the configuration process in a step-wise manner. The tool is coded completely separately from the CMD tool presented in Chapter 2 (§2.5). It is equipped with several GUIs to enhance user-friendliness. A close integration with the InfoWorks CS software was foreseen, although the tool can handle spreadsheet data as well, making it compatible with virtually all software packages or measurement data sets.

To illustrate the presented modelling approach and accompanying software tool, simulation results of a detailed full hydrodynamic InfoWorks CS model were used to set up a conceptual model of the sewer system of part of the city of Ghent in Belgium. The simplified model emulates the CSO results of the detailed model accurately, which was the objective of the study, while the calculation time is reduced by more than 10^6 times compared to the original detailed model.

4

Lumped floodplain modelling using Takagi-Sugeno models

4.1 Introduction

Accurate and fast simulation models of river floodplains play an essential role in many water management applications and decision making, as floodplains exhibit a significant influence on the flow routing and floods are often the subject of the research itself. For instance, inundation depths and flood extents are regularly linked to damage functions (e.g. Jonkman et al., 2008), enabling the optimization of cost-effective water management strategies.

Simulating floods using simplified relationships with computationally efficient models can be very challenging due to the complex river-floodplain interactions, and the mutual interplay between floodplains themselves. This chapter presents an extension to the conceptual modelling approach for rivers presented in Chapter 2 and is specifically aimed at emulating such complex floods. The methodology combines state-of-the-art Adaptive Neuro Fuzzy Inference Systems (ANFIS) of the Takagi-Sugeno type with two different approaches to smooth target values and ensure proper ANFIS training: assuming that the target training values follow a stochastic evolution described by a generalized random walk model, or by simply constraining target values. This combination of techniques results in a very powerful emulation model. The approach can be integrated in the earlier presented modelling approach for rivers and floodplains if desired, but is also compatible with other river models and functions as a separate module.

This research is presented as a separate chapter, since it is not included in the CMD software to conceptualize rivers and floodplains (see §2.5). We deliberately choose not to incorporate this approach in the CMD tool, since it uses advanced and more complex model structures and algorithms that require very specialized expert-knowledge for proper usage. The CMD software, in contrary, is also designed for non-expert users in data-driven modelling. Secondly, the model structures incorporated in the CMD tool succeed in emulating the floods accurately in the vast majority of situations. However, the approach presented in this study provides a powerful tool to emulate more complex flows and floods with enhanced accuracy, and is therefore a valuable extension to the earlier presented methodology.

The next section of this chapter provides a detailed overview of the fuzzy logic techniques used in this study. Section 4.3 presents the methodology to simulate the flow from and to floodplains, and water levels in flooded areas. A simple set of equations is formulated based on water level differences, thereby ensuring that the driving forces of the flow are incorporated in the model. A discharge coefficient is then approximated by an ANFIS model that relies on water levels in the river and/or floodplain. The training values are smoothed before ANFIS training to avoid that the training values diverge to infinity for very small water level differences. This smoothing can be performed by assuming that the coefficients follow a generalized random walk model, or by constraining the target values to a specified range. In addition, a second and completely independent approach is elaborated based on fixed weir equations that are optimized using simulation results of detailed hydrodynamic models. This approach is a simplified version of the methodology presented in Chapter 2 (using weir equations as model structures) and included to compare the use of fixed weir equations with the newly developed approach using ANFIS models. Both approaches were tested on four floodplains along the Dender River. Section 4.4 presents the case study area and available data in detail, while the results are discussed in Section 4.5.

This chapter is based on the following publication:

> WOLFS, V., WILLEMS, P. (2013). A DATA DRIVEN APPROACH USING TAKAGI-SUGENO MODELS FOR COMPUTATIONALLY EFFICIENT LUMPED FLOODPLAIN MODELING. JOURNAL OF HYDROLOGY 503, PP. 222-232.

4.2 Fuzzy Logic

Fuzzy inference systems (FIS) are non-linear modelling approaches that map an input space to an output space using a set of fuzzy "If-Then rules". The classical notion of binary membership in a set has been modified to include partial membership ranging from 0 to 1 (Zadeh, 1965), hereby creating fuzzy sets with imprecise boundaries. Therefore, uncertainties inherent in the system can be reckoned with.

The IF-THEN rule statements are used to formulate the conditional statements that comprise fuzzy logic, e.g. "*IF* antecedent proposition, *THEN* consequent proposition". Depending on the structure of the consequent, the following two types of fuzzy systems can be distinguished: linguistic (Mamdani Type) (Zadeh, 1973; Mamdani 1977) and Takagi-Sugeno (TS) fuzzy models (Takagi and Sugeno, 1985; Sugeno and Kang, 1988). The first group has fuzzy sets in both the antecedent and consequent proposition, whereas TS models have crisp output functions of the input. Considering their simplicity and efficiency, first-order TS models are employed in this work.

4.2.1 Takagi-Sugeno (TS) fuzzy inference system

The process of formulating the mapping from a given input to an output using fuzzy logic is called fuzzy inference (Jang, 1993). Characteristics of input data are allocated to input membership functions, input membership functions to rules, rules to a set of output characteristics and output characteristics to a single-valued output (Jang et al., 2002).

In the first-order TS fuzzy systems, the rule consequent forms a first order polynomial of the input variables $\boldsymbol{x}$:

$$R_i: If\ \ \boldsymbol{x}\ \ is\ A_i\ \ \ Then\ \ \ y_i = \boldsymbol{a}_i^T \cdot \boldsymbol{x} + b_i, \qquad i\ =\ 1, 2, \ldots, M \tag{4.1}$$

where $\boldsymbol{x} \in \Re^p$ is a premise part, $y_i \in \Re$ is the consequent of the ith rule and M is the number of rules. In the consequent, $\boldsymbol{a}_i$ is the parameter vector and b_i the scalar offset. The terms A_i in the antecedents of the rules represent fuzzy sets (Zadeh, 1965) which are used to partition the input space into overlapping regions. Each multivariate fuzzy set A_i of the ith rule is defined by the degree of fulfillment (DOF, μ_i) of the ith rule, which is in this study evaluated using a t-norm (Piegat, 2001; Zimmerman, 2001), such as the algebraic product:

$$\mu_i(\boldsymbol{x}) = \prod_{j=1}^{p} \mu_{ij}(x_j) \tag{4.2}$$

where x_j is the jth input variable in the p dimensional input data space, and μ_{ij} the membership degree of x_j to the fuzzy set describing the jth premise part of the ith rule. This evaluation is also known as the "AND fuzzy operator" with the "product method". Several types of membership functions of the fuzzy sets in the antecedents of the rules are examined later on in this study.

For the input $\boldsymbol{x}$ the total output y of the TS model is computed by aggregating the individual rules contributions:

$$y = \sum_{i=1}^{M} u_i(\boldsymbol{x}) \cdot y_i \tag{4.3}$$

where u_i is the normalized DOF of the antecedent clause of the ith rule:

$$u_i(\boldsymbol{x}) = \frac{\mu_i(\boldsymbol{x})}{\sum_{i'=1}^{M} \mu_{i'}(\boldsymbol{x})} \tag{4.4}$$

4.2.2 Generation of TS fuzzy model

TS fuzzy models can be built using several identification methods. The computation of the antecedent membership functions using Grid Partitioning (GP) is widely used. Alternatively, fuzzy clustering techniques can be employed, which divide the data space into fuzzy clusters. Such methods generate a Fuzzy Inference System (FIS) with the minimum number of rules required to distinguish the fuzzy qualities associated with each of the clusters. Both antecedent identification techniques are described and applied in this study. Next, the consequent parameters can be computed by solving a linear least squares problem.

4.2.2.1 Identification of the antecedent membership functions

a. Grid Partitioning

This method proposes independent partitions of each antecedent variable (Jang, 1993). Using prior knowledge and experience, the modeller can define the membership functions of all antecedent variables. However, the division into the different partitions is often impeded due to the lack of knowledge. Therefore,

the domains of the antecedents are simply divided in homogeneously spaced and equally shaped membership functions. If input-output data is available, the location of these membership functions can be optimized. A major drawback is that the membership functions for every variable are constructed independently of each other. Therefore, relationships between the variables are ignored. Secondly, grid partitioning generates rules by enumerating all possible combinations of membership functions of all inputs. This leads to an exponentially increasing number of rules with increasing number of inputs.

b. Subtractive Clustering

Clustering is a popular unsupervised pattern classification technique which partitions the input space into *C* regions based on similarity or dissimilarity metrics (Jain and Dubes, 1988). Various methods have been described in literature, such as mountain clustering (Yager and Filev, 1994), subtractive clustering (Chiu, 1994) and fuzzy C-means clustering (Bezdek, 1973). In this study the subtractive clustering (SClust) technique is applied, which is an extension of the mountain clustering method.

First, data points are mapped into a unit hyperbox based on their minimum and maximum value. Each data point $\boldsymbol{x}_j \in \Re^p$ of a set of N points is a candidate for cluster centers. A *density measure* D_j is calculated based on the location of the data point with respect to all other data points:

$$D_j = \sum_{k=1}^{N} exp\left(-\frac{\left\|\boldsymbol{x}_j - \boldsymbol{x}_k\right\|^2}{(r_a/2)^2}\right) \tag{4.5}$$

with r_a a positive constant called *cluster radius* and $\|\cdot\|$ the Euclidean distance. The density measure increases when more data points are closer. The data point having the largest density measure, denoted by D_{C1}, is chosen as first cluster center $\boldsymbol{x}_{C1}$. Next, the density measure of all other points is recalculated as follows:

$$D_j = D_j^{original} - D_{C1} \cdot exp\left(-\frac{\left\|\boldsymbol{x}_j - \boldsymbol{x}_{C1}\right\|^2}{(r_b/2)^2}\right) \tag{4.6}$$

with $r_b = \eta \cdot r_a$ the radius defining the neighborhood that will have measurable reductions in density measure, and η a positive constant, denoted as s*quash factor*.

Again, the data point with the largest measure density D_{C2} is taken to be the next cluster center $\boldsymbol{x}_{C2}$, if

$$D_{C2} > \bar{\varepsilon} \cdot D_{C1} \tag{4.7}$$

with $\bar{\varepsilon}$ the *accept ratio*. The data point is still considered as cluster center $\boldsymbol{x}_{C2}$ if the following condition holds, even if the first condition (Eq. 4.7) failed:

$$\frac{d_{min}}{r_a} + \frac{D_{C2}}{D_{C1}} \geq 1 \tag{4.8}$$

with d_{min} the minimal distance between $\boldsymbol{x}_{C2}$ and the previous cluster center. This process will be repeated to obtain new cluster centers. If a potential cluster center does not meet the above described conditions, it is rejected. The density measure of that data point is then set to zero and the one with the next largest density measure is selected and re-tested. The clustering ends if the following condition holds:

$$D_{Ck} < \underline{\varepsilon}\, D_{C1} \tag{4.9}$$

with $\underline{\varepsilon}$ the *reject ratio*. Each cluster center is considered as a fuzzy rule (Chiu, 1994). The DOF μ_i of input $\mathbf{x}$ to rule i is calculated in terms of the distance to the defined cluster centers:

$$\mu_i = exp\left(-\frac{\|\boldsymbol{x} - \boldsymbol{x}_{Ci}\|^2}{(r_a/2)^2}\right) \tag{4.10}$$

4.2.2.2 Computing the consequent parameters

Finally, the linear consequent parameters $\boldsymbol{a}_i$ and b_i of each rule i can be estimated simultaneously solving a least squares problem (Babuška, 1998). For a set of N data points, and a system with M rules, the problem has the following form:

$$\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} = \tag{4.11}$$

$$\begin{bmatrix} u_1(\boldsymbol{x}(1)) \cdot x_1(1) & \cdots & u_1(\boldsymbol{x}(1)) \cdot x_p(1) & u_1(\boldsymbol{x}(1)) & \cdots & u_M(\boldsymbol{x}(1)) \cdot x_1(1) & \cdots & u_M(\boldsymbol{x}(1)) \cdot x_p(1) & u_M(\boldsymbol{x}(1)) \\ \vdots & & \vdots & \vdots & & \vdots & & \vdots & \vdots \\ u_1(\boldsymbol{x}(N)) \cdot x_1(N) & \cdots & u_1(\boldsymbol{x}(N)) \cdot x_p(N) & u_1(\boldsymbol{x}(N)) & \cdots & u_M(\boldsymbol{x}(N)) \cdot x_1(N) & \cdots & u_M(\boldsymbol{x}(N)) \cdot x_p(N) & u_M(\boldsymbol{x}(N)) \end{bmatrix}$$

$$x \; [a_{1,1} \quad \cdots \quad a_{1,p} \quad b_1 \quad \cdots \quad a_{M,1} \quad \cdots \quad a_{M,p} \quad b_M]^T$$

with y_k the outcome of the consequent part for the kth data point, $x_j(k)$ the value of the jth variable in the kth data point, u_i the basis function defined in Eq. 4.4 for $i = 1,\ldots,M$ and $a_{i,j}$ and b_i the consequent parameters defined in Eq. 4.1 for $j = 1,\ldots, p$ with p the dimension of input $\mathbf{x}$.

4.3 Methodology

As described in the introduction, the aim of the proposed methodology is dual: creating a model that can emulate the flow rate between river and floodplain on the one hand and the level in the floodplain on the other hand, based on the water level in the river. The model is derived from time series data of flows and water levels in the river and the floodplain for a number of flood events. The following paragraphs elaborate on the two main goals of the model, namely approximating the flow and stage in the floodplain. Finally, the model's calculation schemes are shown.

4.3.1 Calculation of the flow

In order to have great generalization ability and to enhance interpretability, it is of paramount importance that the model is based on physical principles. The overbank or dike overtopping flow near a floodplain is highly similar to the flow over a weir. The simplest variant of the equations to calculate the flow over a weir is given by:

$$Q = c \cdot b \cdot H^{3/2} \tag{4.12}$$

with c a coefficient of discharge, b the weir opening width and H the head above the crest level. It is endemic to this strong simplification that multiple distinct mechanisms operating over a wide range of time and length scales are not captured accurately (Stewart et al., 1999). The irregular roughness and shape of the embankment for instance will have a significant impact on the calculated discharge. With augmenting level, the width of the overtopping is likely to increase. Additionally, as the stages in the river and the floodplain converge, the free flow can be altered into a reduced drowned flow. To take this latter issue into account, Brater and King (1976) proposed the following factor:

$$Q' = Q\left[1 - \left(\frac{H_2}{H_1}\right)^{1/2}\right]^{0.385} \tag{4.13}$$

with H_1 and H_2 respectively the upstream and downstream head above the weir crest and Q the discharge over the weir that would be obtained in free flow conditions at head H_1. Similar relationships can be found in literature (e.g. Rössert, 1964; Innovyze, 2013).

Furthermore, it is likely that the influencing factors of the flow will show different behaviour for flow from the river to the floodplain and contrariwise. Therefore, the data is divided into two sets based on the stages: one series containing all data with stages in the river surpassing those in the floodplain, and a second series for the opposite case. A model will be built for each of both acquired sets. First, models with a rigid form are discussed. Because of the great number of influencing factors, more flexible and sophisticated models using ANFIS are examined as well.

4.3.1.1 Rigid formulation

The following set of equations was employed in this research, adopted from Innovyze's hydrodynamic software InfoWorks RS, to calculate the flow over certain hydraulic structures (Innovyze, 2013):

$$\begin{cases} Q = 0 & if\ h_{river} \le h_d\ \ and\ \ h_{FP} \le h_d \\ Q = C_1 \cdot (h_{river} - h_d)^{3/2} & if\ h_{river} > h_d\ \ and\ h_{river} > h_{FP}\ \ and\ h_{FP}/h_{river} \le m_1 \\ Q = C_1 \cdot (h_{river} - h_d) \cdot \sqrt{\dfrac{h_{river} - h_{FP}}{1 - m_1}} & if\ h_{river} > h_d\ \ and\ h_{river} > h_{FP}\ \ and\ h_{FP}/h_{river} > m_1 \\ Q = -C_2 \cdot (h_{FP} - h_d)^{3/2} & if\ h_{FP} > h_d\ \ and\ h_{river} < h_{FP}\ \ and\ h_{river}/h_{FP} \le m_2 \\ Q = -C_2 \cdot (h_{FP} - h_d) \cdot \sqrt{\dfrac{h_{FP} - h_{river}}{1 - m_2}} & if\ h_{FP} > h_d\ \ and\ h_{river} < h_{FP}\ \ and\ h_{river}/h_{FP} > m_2 \end{cases} \quad (4.14)$$

with h_r and h_{FP} the water levels in the river and floodplain respectively, h_d the dike level, C_1 and C_2 constant coefficients and m_1 and m_2 modular limits. This type of equations in particular is well-suited to calculate flows over broad crested weirs with rectangular throats, or weirs with a parabolic or triangular control section (Innovyze, 2013). By allocating different coefficients to flow going from the river to the floodplain and vice versa, the equations have greater ability to adapt to the presented time series. The coefficients were optimized using a MATLAB built-in interior-point algorithm (see Byrd et al., 2000). A value of 18.25 m was chosen for h_d, which is equal to the maximum water level in the floodplain or river corresponding to the minimum non-zero flow between river and floodplain.

4.3.1.2 ANFIS approach

The second methodology reasons from a different, more transparent set of equations:

$$\begin{cases} Q = 0 & if\ h_{river} < h_d\ and\ h_{FP} < h_d \\ Q = C_1 \cdot (h_{river} - h_d)^{3/2} & if\ h_{river} > h_d > h_{FP} \\ Q = C_1 \cdot (h_{river} - h_{FP})^{3/2} & if\ h_{river} > h_{FP} > h_d \\ Q = -C_2 \cdot (h_{FP} - h_d)^{3/2} & if\ h_{FP} > h_d > h_{river} \\ Q = -C_2 \cdot (h_{FP} - h_{river})^{3/2} & if\ h_{FP} > h_{river} > h_d \end{cases} \quad (4.15)$$

Again, the second and third equations represent flow from the river to the floodplain, the last two the opposite. To reduce the range of parameters C_1 and C_2, the difference of stages in river and floodplain is used when both surpass the dike level. Both set-ups imply that the latter two parameters should account for all factors affecting the flow, including varying width, roughness, flow conditions (e.g. free versus drowned), etc. Soft computing techniques, such as ANFIS, are due to their flexibility and learning ability likely well qualified for approximating these coefficients. During the model structure identification and training of ANFIS, target values must be presented. These values can be obtained by simply dividing the observed flow rates by the predicted flow rates (Eq. 4.15) for all n time steps:

$$C_i = \frac{Q_{observed,i}}{Q_{predicted,i}} \quad for\ i = 1, 2, \dots, n \quad (4.16)$$

However, this simple approach induces values striving to infinity for calculated flows around zero. Such target values encumber proper ANFIS training. Therefore, a pre-processing procedure is compulsory. A plausible approach is the assumption that the coefficients C show a stochastic evolution described by a generalized random walk process (Jakeman and Young, 1981), such as a Random Walk, Integrated Random Walk, autoregressive process, Smoothed Random Walk, etc. Hence, the sought coefficient is a time varying parameter that can be estimated via the following dynamic linear regression:

$$Q_{observed,i} = C_i \cdot Q_{calculated,i} + e_i \quad for\ i = 1, 2, \dots, n \quad (4.17)$$

where e_i is an irregular component, normally defined for analytical convenience as a serially uncorrelated and normally distributed Gaussian sequence with zero

mean. The calculated flow obtained via Eq. 4.15 acts as an exogenous regression parameter. The Captain Toolbox (Taylor et al., 2007; Young et al., 2007) for MATLAB provides various functions for extensive time varying parameter analyses (e.g. Young et al., 2001). Depending on the chosen generalized random walk process, the acquired coefficients will follow a smoother course, obviating leaps towards infinity. However, incautious process association or improper use of the available tools can deliver erroneous estimations and impose trends to the target values of the ANFIS, yielding inaccurate results. For that reason, this study advocates a much simpler yet efficient approach, by constraining the target values obtained via Eq. 4.16: all values surpassing a certain threshold (e.g. the median of the retrieved values multiplied by a chosen factor) are removed prior to ANFIS identification and training. This ensures that the ANFIS construction focuses on most data points instead of some outliers. In addition, rescaling the input-output space to a unit hyperbox is possible as often proposed (e.g. Vernieuwe et al., 2006).

Two ANFIS models will be built to approximate C_1 and C_2 separately, denoted as 'ANFIS1' and 'ANFIS2' respectively. The most relevant inputs for the ANFIS are unknown, but it seems reasonable to relate the coefficients C to the stages in the river and floodplain. Therefore, ANFIS are built with two different input sets: ANFIS1 with one input (h_{river}-h_d) and two inputs (h_{river}-h_d and (h_{FP}-h_d)/(h_{river}-h_d)), and ANFIS2 with one input (h_{FP}-h_d) and two inputs (h_{FP}-h_d and (h_{river}-h_d)/(h_{FP}-h_d)). This set-up is closely related to the aforementioned theory (see Eq. 4.13), but allows for capturing other dynamics than submerged flow as well. The two identification techniques discussed earlier, namely grid partitioning and subtractive clustering are examined. The obtained ANFIS will be trained via a hybrid learning algorithm.

4.3.2 Calculation of the stage in the floodplain

The water level in the floodplain is calculated via a stage-volume relationship (denoted as hypsometric curve; see also §2.3.3.1). The hypsometric curve itself is determined via retrieved stage-volume couples from simulation results and parameterized by means of a simple univocal piecewise-linear relationship.

For larger floods, one can follow a different approach to increase accuracy and robustness. When the water level in the river exceeds a certain threshold, the water level in an adjacent floodplain often becomes linearly correlated or even equal to that of the river. The floodplains and river seem to merge into one

channel. Figure 4.1 shows water levels of the Dender river (h_{river}) and a neighbouring floodplain (h_{FP}) acquired via simulations with a full hydrodynamic model. One can clearly notice the univocal relationship between both water levels (further denoted in the text as an 'H-H relationship') that surpass 18.5 m. Hence, for larger floods, the water level in the floodplain can directly be linked to that in the river, without the need of calculating the flow rate. However, knowledge of the discharge between river and floodplain at all times is important for most hydrodynamic river models. For such cases, the discharge can be estimated via the inverse hypsometric curve and the volume at the previous time step.

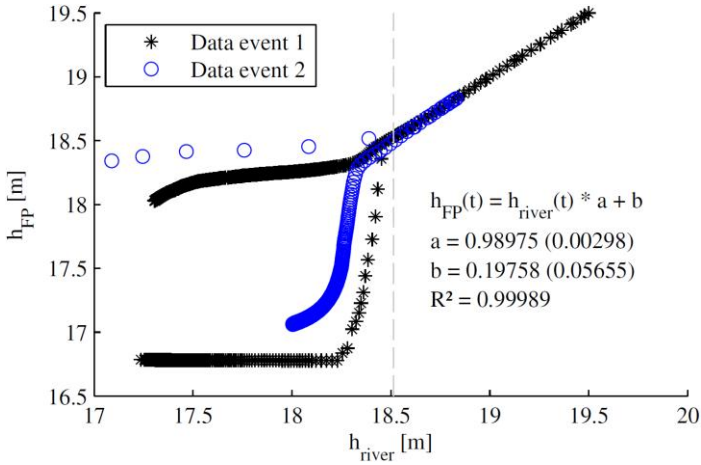


Figure 4.1: Water levels of the river (h_{river}) and contiguous floodplain (h_{FP}) obtained using simulations with a full hydrodynamic model. The parameters of the linear regression curve (H-H relationship) are shown for river water levels exceeding 18.5 m, together with their 95% confidence intervals.

4.3.3 Calculation schemes

The above methodology leads to two different calculation schemes, shown schematically in Figure 4.2. For each time step, one of the two schemes is executed. Note that the second calculation scheme can only be followed if the stages in river and floodplain exhibit an (approximate) univocal relationship for certain stages (e.g. Figure 4.1).

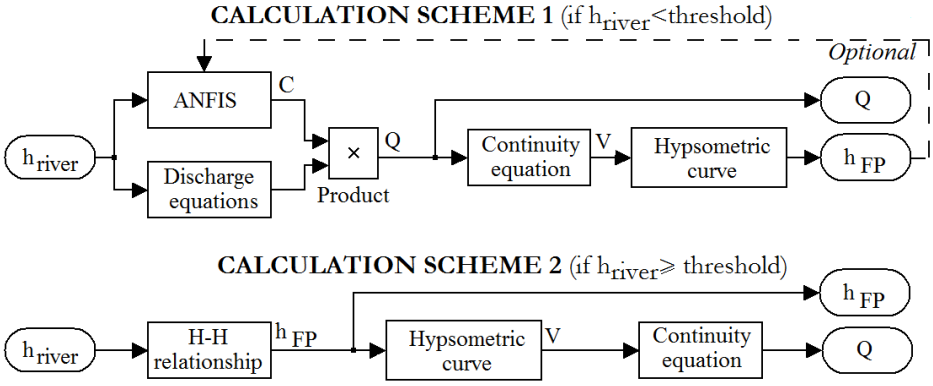


Figure 4.2: Schematic overview of the two calculation schemes.

The water level in the river h_{river} comes in this study from a full hydrodynamic MIKE model. In practice, it can be a measured water level, or a simulated value from another model that is likely affected by the flow between the river and the considered floodplain. If the stage in the river does not exceed the specified threshold value, the first scheme is followed. In this case, Eq. 4.14 or 4.15 are used to calculate the discharge. Next, the discrete continuity equation is used to calculate the volume in the floodplain, which in turn leads to a water level estimate via the hypsometric curve. Scheme 2 on the other hand determines the stage in the floodplain directly through the H-H relationship. If desired, the flow between the river and the floodplain can be rated via the continuity equation.

4.4 Study area

The modelling approach will be tested on four floodplains (denoted as 'FP 1 – 4') located along the navigable river Dender in Belgium. The four floodplains are situated near the city of Geraardsbergen. Their sizes vary between 4 and 104 ha. Floods ravage the area regularly: larger historical floods occurred in 1993, 1995, 1999, 2002 and 2010. Whereas the average discharge at the upstream gauging station of Overboelare is 5 m^3/s, the discharge rose to more than 100 m^3/s during the flood of November 2010.

Due to the scarcity of accurate measurement data, simulation results from a detailed, one-dimensional full hydrodynamic MIKE11 model were used. Floodplains were incorporated in a quasi 2D setting to reduce calculation costs (Willems et al., 2002; Willems, 2013a): they are described as a network of

fictitious river branches and spills connected to the main river. The model was built using two dimensional spatial information of the floodplain's topography, and measurements of water levels and flows (Figure 4.3).

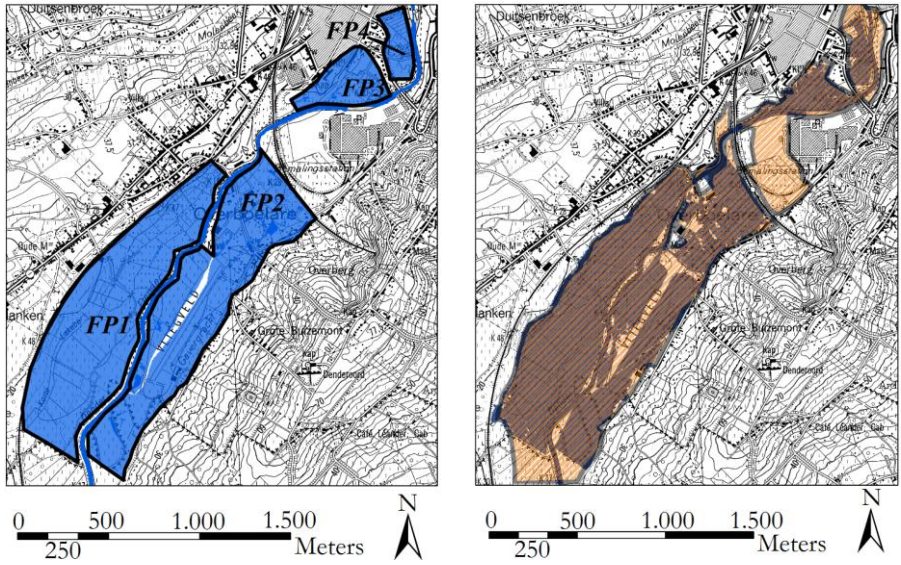


Figure 4.3: (Left) Topographic map of the city of Geraardsbergen and the Dender river. The cross sections of the hydrodynamic model of the four modelled floodplains (FP 1 – 4) cover the shaded areas. (Right) The flooded are of the storm of November 2010 (dashed orange) versus the modelled flooded area (area shaded in blue).

After calibration and validation of the MIKE11 model for historical flood periods, simulations were also conducted for synthetic runoff events. These represent input hydrographs for given return periods. See Vaes et al. (2000), Willems et al. (2002) and Willems (2013) for more details on these synthetic events and the simulations.

Simulation results of three events, each lasting around one month, were employed in this study. The results were saved on a 10-minute base, resulting in 4320 time samples for each event. The data used to identify and train the ANFIS consists of the MIKE11 synthetic event simulation results for a return period of 100 years (denoted as 'event1'). This event led to a flood in all four floodplains. The reason for selecting a complete event as input is to include all different overflow conditions, which should result in a more robust model. Simulation results of the historical storm of November 2010 ('event2') were used as checking

data to prevent overfitting of the ANFIS, while results of the synthetic design storm with a return period of 25 years ('event3') were employed as (independent) testing data. In this study, the acquired simulation results of stages along the river and the floodplain are averaged over the calculation nodes for simplicity, so only one stage in river and one in the floodplain are retained. Accordingly, the discharges over all spills between the river and the floodplain are cumulated, resulting in one time series of flow data.

4.5 Results and discussion

The model performance is evaluated by means of the following indices:

(i) The criterion of Nash and Sutcliffe (1970) (NSE): see Eq. 2.7

(ii) The root mean square error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_{observed,i} - Y_{predicted,i})^2}{n}} \tag{4.18}$$

(iii) The bias (B):

$$B = \frac{\bar{Y}_{predicted}}{\bar{Y}_{observed}} \tag{4.19}$$

(iv) The mean absolute error (MAE):

$$MAE = \frac{\sum_{i=1}^{n}\left|Y_{observed,i} - Y_{predicted,i}\right|}{n} \tag{4.20}$$

with $Y_{observed,i}$ the value of the full hydrodynamic simulation result of parameter Y for time i, $Y_{predicted,i}$ the outcome of the model for time i, $\bar{Y}$ the average of parameter Y over all time steps and n the number of time steps. Only data points are considered of time steps with $Q_{observed}$ different from zero, hence the beginning and ending of the simulation, when there is no flooding, are not taken into account.

4.5.1 Evaluation rigid formulation approach

First, the performance of Eq. 4.14 is investigated for modelling FP1. Optimization with the data of event 1 yielded values of 230.62 and 35.87 for C_1 and C_2 respectively, and 0.95 and 0.57 for m_1 and m_2. The statistical performances of the flow and the stage in the floodplain are summarized in Table 4.1. The top three rows of the table ('Using observed h_{FP}') represent a model that does not use the calculated floodplain stages as input in following time steps, but apply instead the observed floodplain stages from the hydrodynamic simulation results. This simulation allows to assess the model performance without the propagation of errors. The last three rows demonstrate model performance in "full simulation mode", meaning that the model employs previously calculated floodplain water levels as inputs for subsequent time steps. At each time step, the model follows calculation scheme 1 (see Figure 4.2) using a piecewise linear hypsometric curve.

Table 4.1: Statistical performances of Eq. 4.14 for the calibration event (a) and the two validation events (b and c) for FP1 when using observed and calculated (i.e., full simulation) values of the stage in the floodplain as input.

	Flow			**FP stage (h_{FP})**
	NSE [-]	RMSE [m^3/s]	Bias [-]	MAE [10^{-3} m]
Using observed h_{FP}	(a) 0.644	(a) 3.980	(a) 1.271	(a) n/a
	(b) 0.934	(b) 1.051	(b) 1.070	(b) n/a
	(c) 0.707	(c) 4.202	(c) 0.886	(c) n/a
Using calculated h_{FP}	(a) -3.445	(a) 14.072	(a) 1.073	(a) 22.2
	(b) 0.923	(b) 1.129	(b) 1.031	(b) 20.3
	(c) 0.389	(c) 6.064	(c) 1.043	(c) 22.0

The model succeeds only moderately in emulating the flow series when presenting observed floodplain stages. When the model relies on self calculated floodplain levels however, the accuracy drops dramatically for two out of the three investigated events. It is clear that the carefully chosen rigid functional form does not suffice for approximating the flow. Therefore, the use of more complex methods and model structures is necessary.

4.5.2 Evaluation ANFIS approach

4.5.2.1 Floodplain 1

Two different ANFIS identification techniques, namely grid partitioning (GP) and subtractive clustering (SClust), both discussed in §4.2.2.1, are examined by modelling floodplain 'FP 1'. Moreover, two input sets are considered to investigate in a model-based ad-hoc approach which tackles the most uncertainties entailed by the presented modelling methodology. Depending on the results, a selected identification technique and input set will be utilized to test model performance on the data of FP 2 – 4. For larger floods, an H-H relationship can be incorporated as well. This relationship is shown in Figure 4.1 for FP1. The same hypsometric curve was employed as in the evaluation of the rigid form approach (§4.5.1) in all cases. This led to 8 different models, summarized in Table 4.2.

First, the fuzzy rule-base determination of the TS model and membership function initialization was done using the GP method. Models were automatically built with the following membership function (MF) types: triangular ('trimf'), trapezoidal ('trapmf'), generalized bell curve ('gbellmf'), Gaussian curve ('gaussmf') and two-sided Gaussian curve ('gauss2mf'). The number of MFs was varied between 2 and 5 for each of the antecedent variables. Their parameters were optimized on the training data set through back propagation. The training set contains the data from event 1 and comprises 65 and 297 points for ANFIS 1 and 2 respectively. The parameters of the consequent parts were calculated using a linear least squares method (see §4.2.2.2). The training epoch number for the optimization was set to 100. As indicated earlier, event 2 was used as checking data to prevent overfitting. The checking set contained 67 and 264 points for ANFIS 1 and 2 respectively. The ANFIS structure was chosen according to a minimum checking error criterion. Of all examined models, the one with lowest RMSE of training and checking data was retained. The preserved structures and the number of parameters are added to Table 4.2.

The ANFIS identification and parameter computation was also done using the SClust method. In order to find the most appropriate model, the cluster radius r_a was varied between 0.1 and 1 with steps of 0.05. The squash factor, accept ratio and reject ratio were kept constant with values of 1.25, 0.5, and 0.15 respectively. The same procedure as for the GP model was performed for the optimization.

Again, the model with lowest RMSE of training and checking data was preserved. Their configurations are appended to Table 4.2.

Table 4.2: Properties of the 8 retained models for ANFIS1 and ANFIS2 for FP1 with H_{river} and H_{FP} the head above the dike, h_{river} and h_{FP} the water level.

	inputs ANFIS1; ANFIS2	GP /SClust	H-H module	#rules ANFIS1; ANFIS2	MF Type	#parameters ANFIS1; ANFIS2
Model 1-a	H_{river}; H_{FP}	GP	Incl.	5; 5	gbellmf	25; 25
Model 1-b	H_{river}; H_{FP}	GP	Excl.	4; 4	gaussmf	16; 16
Model 2-a	H_{river}; H_{FP}	SClust	Incl.	7; 2	n/a	28; 8
Model 2-b	H_{river}; H_{FP}	SClust	Excl.	5; 3	n/a	20; 12
Model 3-a	H_{river}, h_{FP}/h_{river}; H_{FP}, h_{river}/h_{FP}	GP	Incl.	4; 25	gbellmf	24; 105
Model 3-b	H_{river}, h_{FP}/h_{river}; H_{FP}, h_{river}/h_{FP}	GP	Excl.	4; 16	gbellmf	24; 72
Model 4-a	H_{river}, h_{FP}/h_{river}; H_{FP}, h_{river}/h_{FP}	SClust	Incl.	3; 3	n/a	21; 21
Model 4-b	H_{river}, h_{FP}/h_{river}; H_{FP}, h_{river}/h_{FP}	SClust	Excl.	3; 4	n/a	21; 28

To limit the number of required experiments, the same MF types were used for ANFIS1 and ANFIS2. The MF types of the models with best performance were retained, but it should be noted that the differences in accuracy among the tested MF types were small. Table 4.3 shows the values of the performance indices and error functions for the calibration (a), checking (b) and testing (c) events. Only the MAE is shown for the stage estimations, because all NSE indices of the floodplain levels lay between 0.97 and 1. It is important to stress that these values were achieved in full simulation mode, i.e., the model uses self calculated floodplain stages as input.

Table 4.3: Statistical performances of the ANFIS approach for the training (a), checking (b) and testing (c) events for FP1 in full simulation mode.

	Flow			**FP stage (h_{FP})**
	NSE [-]	RMSE [m^3/s]	Bias [-]	MAE [10^{-3} m]
Model 1-a	(a) 0.963 (b) 0.972 (c) 0.912	(a) 1.290 (b) 0.522 (c) 1.761	(a) 0.986 (b) 1.005 (c) 0.820	(a) 4.4 (b) 5.2 (c) 3.3
Model 1-b	(a) 0.949 (b) 0.971 (c) 0.926	(a) 1.516 (b) 0.530 (c) 1.614	(a) 1.027 (b) 1.005 (c) 1.018	(a) 12.1 (b) 5.3 (c) 5.4
Model 2-a	(a) 0.963 (b) 0.972 (c) 0.913	(a) 1.296 (b) 0.519 (c) 1.749	(a) 0.985 (b) 1.004 (c) 0.821	(a) 4.8 (b) 5.5 (c) 3.6
Model 2-b	(a) 0.946 (b) 0.971 (c) 0.925	(a) 1.565 (b) 0.533 (c) 1.621	(a) 1.035 (b) 1.012 (c) 1.025	(a) 11.8 (b) 4.6 (c) 4.6
Model 3-a	(a) 0.972 (b) 0.961 (c) 0.895	(a) 1.120 (b) 0.612 (c) 1.920	(a) 0.986 (b) 1.004 (c) 1.074	(a) 5.7 (b) 6.0 (c) 23.8
Model 3-b	(a) 0.958 (b) 0.961 (c) 0.919	(a) 1.371 (b) 0.616 (c) 1.687	(a) 1.020 (b) 1.054 (c) 1.070	(a) 13.9 (b) 6.6 (c) 6.3
Model 4-a	(a) 0.970 (b) 0.962 (c) 0.911	(a) 1.160 (b) 0.610 (c) 1.774	(a) 0.988 (b) 1.039 (c) 0.834	(a) 5.2 (b) 6.0 (c) 3.6
Model 4-b	(a) 0.956 (b) 0.961 (c) 0.921	(a) 1.412 (b) 0.616 (c) 1.669	(a) 1.034 (b) 0.997 (c) 1.009	(a) 12.3 (b) 7.4 (c) 7.2

As shown in Table 4.3, the models using ANFIS perform much better than the investigated fixed equations (see Table 4.1 and Figure 4.4), with practically all NSE values of flow prediction above 90%. It is clear that the models using soft computing techniques outperform the model based on the rigid functional form. Although the differences in performance among the ANFIS based models are small, it is interesting to highlight some elements.

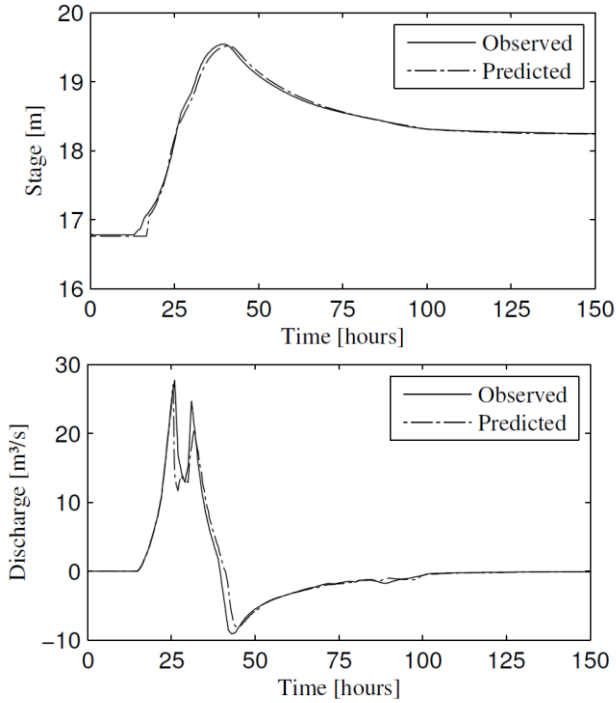


Figure 4.4: Conceptual (dash dotted line) versus full hydrodynamic MIKE11 (full line) model results of stages (left) and discharges (right) for part of the calibration event. The utilized conceptual model uses ANFIS with 1 input each derived using the GP technique (model 1-b).

Models incorporating H-H relationships manage to emulate the stages in the floodplain better than the models that use solely calculation scheme 1, except for the prediction of the testing event by 'Model 3-a'. In that particular case, it is likely that the model has been overfitted, considering the amount of parameters involved (see Table 4.2). The differences in the flow rate predictions between models with and without H-H relationships are rather small, apart from the bias. The flow approximations of the validation event given by the models with H-H module seem in most cases a bit biased. Hence, when obtaining accurate stage

predictions is the primary goal and the discharge values are accessory, models that incorporate H-H relationships are preferable.

Additionally, some differences can be noticed as another ANFIS construction method is employed. Construction of the fuzzy rule base using GP and SClust delivers generally similar results in this test. The advantage of clustering techniques is that the generated rules are more tailored to the input data than they are in a FIS generated without clustering. Hereby, the problem of combinatorial explosion of the rules is reduced when the input data has a high dimension. This is also reflected by the retained model structures listed in Table 4.2: the models built using the SClust technique contain in general fewer rules. However, since the dimension of the input is only one or two in this setting, the number of rules will likely not cause insuperable difficulties. On the other hand, the GP technique can apply different types of membership functions, whereas the degree of fulfilment in SClust is always calculated in terms of the distance to the defined cluster centers (see Eq. 4.10). This reduces the flexibility of the ANFIS somewhat. The differences in accuracy between models built using GP and SClust are negligible in this study though. Hence, both antecedent identification techniques are equally suitable in this case. Because the selection of appropriate cluster centers requires close examination of the inputs and obtained results, the easier to use GP technique is applied when modelling floodplains 2 to 4.

Finally, a distinction can be made based on the basis of the chosen input set. Incorporating multiple inputs augments model complexity, but also increases model flexibility and prediction potential. Figure 4.5a shows the (normalized) outcome of ANFIS1 obtained via the SClust technique in function of the inputs H_{river} and h_{FP}/h_{river}. The outcome of ANFIS1 is factor C_1 which is used to correct the primary discharge estimation in Eq. 4.15. Figure 4.5b represents two cross sections of the surface plot on the left, which show the normalized and averaged output in function of H_{river} and h_{FP}/h_{river} respectively. As expected, the multiplication factor C_1 is a monotonously increasing function of H_{river}. This relationship resembles for instance the increasing width of the overflow with augmenting water level in the river. On the other hand, C_1 will decrease with h_{FP}/h_{river} striving to unity. This reflects the increasing damping of the flow rate as both water levels converge, similar to for instance the submerged character of flows. Hence, the used models in combination with appropriate input sets succeed in generating meaningful ANFIS. The performance of all models using ANFIS with two inputs surpasses that of their simpler counterparts for the calibration event (e.g. 'Model 1-a' versus 'Model 3-a', etc.). For the validation

events however, the models seem to be more inaccurate. As stated earlier, this is likely caused by the higher number of parameters that require calibration in these more complex models. Thus, when using solely one event as training data, the models are somewhat overparameterized. Therefore, the use of only one input for each ANFIS is recommended when calibration data is scarce.

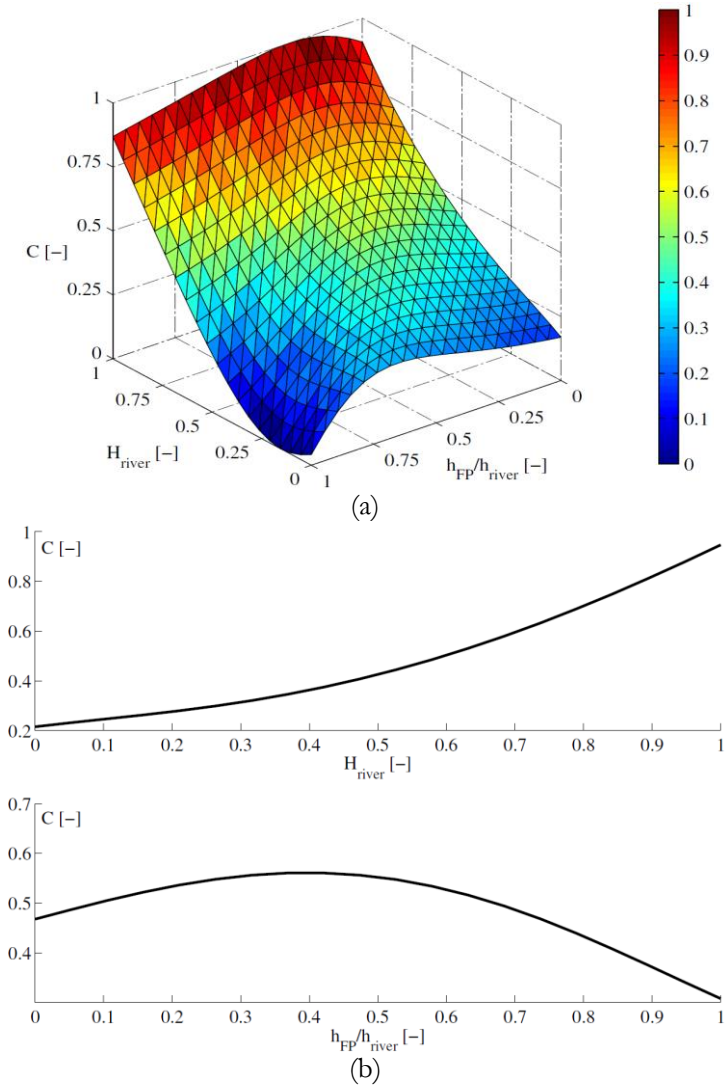


Figure 4.5: Three dimensional normalized outcome of ANFIS1 created via the SClust technique in functions of its two normalized inputs, with H_{river} and H_{FP} the head above the dike, h_{river} and h_{FP} the water level (a). Two orthogonal cross sections with averaged values of the surface plot (b).

4.5.2.2 Floodplains 2-4

To check the robustness and accuracy of the chosen modelling approach, three other floodplains are modelled. Only one input is used for each ANFIS model to reduce the risk of overfitting. The GP technique is employed to identify the rule base and antecedent variables. The calibration procedure is the same as earlier. No flooding occurred in floodplains 3 and 4 during the second event. To provide checking data and obviate overfitting in these cases, the data of event 1 was divided in a randomly unsupervised manner into 70% training and 30% checking data. Such random division is possible since the model structure is not autoregressive and the inputs do not rely on consecutive (differences of) system states. Again, a univocal relationship between the stages in the river and the floodplain is used above a certain threshold value, meaning that calculation schemes 1 and 2 are employed. Model properties are summarized in Table 4.4, together with the results for the examined events.

Table 4.4: Properties of the retained models and the statistical performances of the ANFIS approach for the training (a), checking (b) and testing (c) events for floodplains 2 to 4 (FP 2-4) in full simulation mode.

	# rules ANFIS1; ANFIS2	MF Type	**Flow**			**FP stage (h_{FP})**
			NSE [-]	RMSE [m^3/s]	Bias [-]	MAE [10^{-3} m]
			(a) 0.943	(a) 0.975	(a) 0.806	(a) 16.4
FP 2	3; 2	Gaussmf	(b) 0.452	(b) 1.02	(b) 0.958	(b) 12.5
			(c) 0.684	(c) 1.484	(c) 0.935	(c) 7.3
			(a) 0.973	(a) 0.337	(a) 0.963	(a) 1.5
FP 3	2; 3	Gaussmf	(b) n/a	(b) n/a	(b) n/a	(b) n/a
			(c) 0.922	(c) 0.389	(c) 0.987	(c) 1.6
			(a) 0.966	(a) 0.152	(a) 0.703	(a) 1.8
FP 4	5; 2	Gaussmf	(b) n/a	(b) n/a	(b) n/a	(b) n/a
			(c) 0.966	(c) 0.094	(c) 1.044	(c) 1.8

Again, the lumped modelling approach manages to mimic the full hydrodynamic model adequately. The predicted floodplain stages show in general very low errors. NSE values for the calculated flow series are close to unity for the

calibration events, while the efficiency dropped only little for the validation events. The model of the second floodplain (FP 2) seems to have difficulties to emulate the results for the second and third events though. Close inspection of the retained models is advised, in particular for ANFIS models due to their great flexibility. Hence, simply selecting the model with a best performance statistic could lead to poor predictions for other events.

4.5.3 Computational efficiency

Finally, the calculation speeds are compared of the presented methodology and the original full hydrodynamic implementation in MIKE11. Calculation scheme 1 was entirely implemented in the programming language C. Of course, determining the outcome of ANFIS is a computationally demanding issue. The examined systems merely have one or two inputs and therefore, it is expedient to create a one or two dimensional lookup table. This table is filled prior to simulation, containing the predicted discharges obtained via calculation scheme 1. During a simulation, intermediate values are calculated via linear interpolation. The finer the mesh of the table, the closer the predictions will be to the ANFIS model. A dense grid means an increased computation time however. In this experiment, a two dimensional table was created that uses the stages in the river and floodplain as inputs with a spatial resolution of 0.005 m. The maximum heads in the river and the floodplain above the dike of the training event were extrapolated 15% to cope with values outside the training space, although already a storm with a return period of 100 years was used during training. This dense resolution resulted in a 300 by 600 grid. Simulating the first event via calculation scheme 1 took 0.017 seconds (averaged over 10 runs) on a computer equipped with an i3 processor using a single core, including calling and saving costs. It is of course to be expected that if the computationally less demanding calculation scheme 2 was also adopted, the calculation time would even drop further. The simulation of event 1 in the MIKE11 model with a time step of 24 seconds on the same computer took approximately 41 minutes. The entire model consists of 3013 calculation nodes. The floodplains are incorporated in a computationally efficient quasi-2D setting (Willems et al., 2002). In this case, FP 1 is described by 55 calculation nodes. Assuming that the total calculation time of the model is directly proportional to the number of calculation nodes, predicting the floodplain levels and discharge takes 44.91 seconds. Hence, the conceptual model is more than 2500 times faster than the detailed hydrodynamic model in this

experiment. Note that this computational gain is case specific and will differ for other setups.

4.6 Conclusions

This chapter presented a lumped data-based modelling approach to calculate the water level in a floodplain and the flow rate between river and floodplain, emanating from the level in the river. The developed approach is an extension to the conceptual modelling methodology presented in Chapter 2 and is not included in the CMD software given its complexity and since the CMD software can already simulate the vast majority of floods with sufficient accuracy. The methodology presented in this chapter can be highly effective though when dealing with complex floodplain interactions, or very high simulation accuracies are required.

First, a fixed set of general weir equations is used to predict the flow rate based on the levels in the river and the floodplain. Next, a slightly modified form of the equations is examined in combination with a dynamic multiplication factor which should tackle the altering response characteristics of the system. Two adaptive neuro fuzzy inference systems (ANFIS) are employed with different input sets and build-up methods (grid partitioning (GP) and subtractive clustering (SClust)) to parameterize the dynamic parameter. Neuro-fuzzy systems have become popular in water engineering due to their capability for identifying complexities underlying non-linear systems. However, they do not exhibit a clear declarative knowledge structure. The proposed methodology has an architecture which combines the benefits of the neuro-fuzzy modelling structure with a priori knowledge of the underlying processes. Still, attention must be paid during model set-up to avoid overfitting. Finally, levels in the floodplain are derived from stage-volume relationships, or directly from the water level in the river. This led to two different calculation schemes.

The developed methodology was tested on four floodplains along the river Dender in Belgium. Due to the lack of accurate measurement data, simulation results of three events from a detailed full hydrodynamic model were used. The rigid equations manage to mimic the target flow values adequately during calibration, but fail entirely when previously self-calculated values of floodplain stages are reemployed. In contrast, several performance criteria indicated that the modelling approach using ANFIS delivered precise results: the Nash-Sutcliffe

efficiency values are very close to unity for all calibration events, while they only dropped slightly during validation. The results of models built with GP and SClust performed almost equally, although the GP technique is preferred because of the simpler calibration process. Solely the incorporation of one or two input variables suffices, facilitating the interpretation of the obtained system.

The presented methodology predicts discharges and stages in the floodplain more than 2500 times faster than the full hydrodynamic model for the examined case study. Although the spatial resolution of the water level estimations is limited, the developed methodology has numerous direct applications. Using the presented methodology, floodplains can easily be incorporated explicitly in other models describing the flow in the river.

5

Modelling hysteretic discharge-stage curves

5.1 Introduction

Rating curves, which relate a water level in a river to a discharge via some predefined relationship, play an eminent role in water management. Simple univocal relationships between the water level and discharge rarely exist in reality due to various effects. Most often, these relationships are characterized by hysteretic behaviour, meaning that they have a non-univocal or looped relationship. This problem proved to be an interesting topic to apply, examine and compare different model structures. The research on capturing hysteretic behaviour has great potential for modelling rating curves as demonstrated in this chapter, but the investigated structures were subsequently also incorporated in the general conceptual modelling approach for rivers and sewers presented in Chapters 2 and 3. Their use is not limited to solely simulating flows based on water levels (as in rating curves), but the structures were employed in various settings.

This chapter focuses on several methods for discharge estimation based on time-series of stage data, while accounting for the hysteretic behaviour of the relationship. Since usually in practice only stage-discharge couples are available to set up rating curves, this research concentrates on the use of rating curves that rely solely on this information. Hence, approaches are investigated that try to emulate looped rating curves when twin stage or velocity measurements are

unavailable. To restrain the problem's complexity and obviate measurement uncertainties, only data is used from numerical simulations with models that solve the complete de Saint-Venant equations. Two fundamentally different case studies are examined to assess the performance of the selected modelling methodologies. The stage and discharge data employed for the first case study are gathered at a specific location where variable channel storage and significant backwater result in a large looped rating curve. The rating curve of the second case study is characterized by numerous smaller hystereses, caused by tidal influences downstream. This location is also affected by nearby floodplains during floods. Other influencing factors such as vegetation, sedimentation and erosion are not incorporated in the detailed hydrodynamic models used to generate the case study data and are therefore not taken into account in this research.

First, an introduction on the importance and use of rating curves is given, together with a synopsis of different approaches to model rating curves. Section 5.3 extensively discusses the investigated case study areas and available data. Next, Section 5.4 presents four different approaches to model rating curves. The most commonly used single rating curve approach is discussed, followed by a variant of the Jones' formula in which the conventional rating curve is corrected by a time varying parameter to acquire greater flexibility. This should allow the model to account for the unsteadiness, backwater and variable channel storage effects. A state dependent parameter (SDP) algorithm is used for the non-parametric identification of the dynamic parameter. In addition, two expert systems are discussed, namely the rarely used M5' model tree and artificial neural networks. Section 5.5 examines the results of each modelling approach for both case studies. Section 5.6 investigates the model's sensitivity to small data sets for configuration, discontinuities or outliers in the calibration data, and how each structure behaves during extrapolation.

This chapter is based on the following publication:

> WOLFS, V., WILLEMS, P. (2014). DEVELOPMENT OF DISCHARGE-STAGE CURVES AFFECTED BY HYSTERESIS USING TIME VARYING MODELS, MODEL TREE AND NEURAL NETWORKS. ENVIRONMENTAL MODELLING & SOFTWARE, 55, 107-119.

5.2 Problem statement

Discharge-stage curves, also referred to as rating curves, play an essential role in hydrology. It is common practice to obtain discharge estimates using stage measurements in combination with river stage-discharge relationships. These stream flow data form, in turn, a key source of information in virtually all hydrological applications, such as calibration and validation of rainfall-runoff models, boundary conditions of flood inundation models, stochastic modelling of river flow time series, river sediment studies, etc. Another application is the use of conceptual models that try to emulate results of detailed hydrodynamic models. In such case, the stage-discharge relationship is utilized in the reverse way than described above for the common use of these relationships.

Modelling stage-discharge relationships is difficult due to the large number of influencing factors. Boyer (1964) composed a conclusive list of factors affecting a rating curve:

- Backwater effects – changes in the downstream conditions such as the effects of control constructions and confluence of downstream tributaries. Note that constant backwater, as caused by rigid section controls for instance, will not affect the simple stage-discharge relation detrimentally (Herschy, 1995);
- Unsteadiness of the river flow;
- Variable channel storage – overflow streams onto floodplains during high discharges, thereby resulting in different surface slopes and unsteadiness effects;
- Channel modifications due to dredging, construction works, etc;
- Sediment transport – growing and receding bed forms during floods alter bed roughness, which causes looped rating curves (see e.g. Simons and Richardson, 1962). In addition, sedimentation and erosion can change the cross-section and bed slope of the channel, hereby also affecting the rating curve;
- Vegetation – effect on the roughness and hence the stage-discharge relationship;
- Ice.

These factors can result in rating curves with looped trajectories, denoted as hysteretic behaviour. During flood events for instance, the water surface slope

will depend on whether the discharge is increasing or decreasing. As the discharge increases, the surface slope becomes greater than the slope for steady flow at the same stage. Hence, the discharge in the river is greater than the steady rating curve would suggest. This is also illustrated in Figure 5.1: a measured stage $S_{measured}$ yields via a univocal rating curve a discharge value Q_{calc} that differs from the real values during a flood event ($Q1$ and $Q2$ in the rising and falling limbs respectively). Note that the size and form of a hysteresis is different for each flood (Fread, 1975). More details on the physical background of hystereses can be found in literature (e.g. Chow, 1959; Henderson, 1966; Fenton and Keller, 2001).

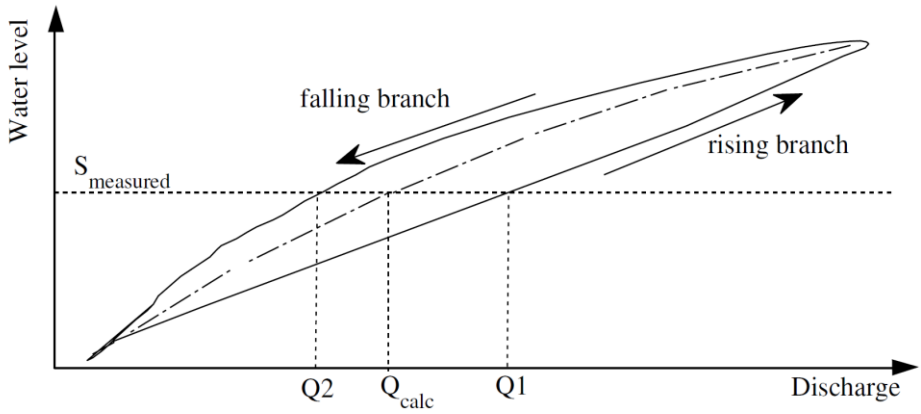


Figure 5.1: Discharge-stage data couples affected by unsteadiness effects showing a looped trajectory (full line) versus the steady rating curve (dash dotted line).

Since hystereses can have a significant influence on the rating curve, it is essential to include this behaviour in the mathematical models describing the stage-discharge relationship. Logically, all hydrologic investigations and practical operations that rely on rating curves can be affected by these hystereses. During highly dynamic floods for instance, the peak discharge could be significantly under- or overestimated. Di Baldassarre and Montanari (2009) showed in a numerical study on the River Po that the errors in discharge estimation when significant flood waves occur may exceed 15% when a flow rating curve is employed that does not take hysteresis into account. Moreover, the arrival time of the peak discharge could be in error and hence influence flood warning predictions, since maximum stage and maximum discharge do not necessarily coincide (Dottori et al., 2009; Fenton and Keller, 2001). Similarly, the calibration of unsteady flow models can be distorted because of incorrect discharge estimates obtained via univocal rating curves.

Numerous approaches can be found in literature that deal with rating curves, each with its own advantages and limitations. The simplest and at the same time most commonly used approach is the single (steady) curve methodology. This approach links the measured water level univocally to a discharge, thereby neglecting effects such as unsteadiness and backwater. The widely known rising and falling curves approach is highly similar to the single rating curve. Prior to the curve estimation, the data set is divided into two groups. One group represents the data for the rising branch, the other the data of the falling branch. By calibrating different curves to both data sets, hysteretic behaviour can be captured to some extent in contrast to the single rating curve approach. However in practice, the classification into the two sets is subjective and not always straightforward. Secondly, discharge estimates obtained via this approach can show sudden drops and rises when shifting between the rising and falling curves (Tawfik et al., 1997). To overcome this issue, most standard hydrometric literature (e.g. Boyer, 1964; Mander, 1978; Herschy, 1995; ISO, 1998) recommends the use of Jones' formula (Jones, 1916) to correct the single rating curve when unsteadiness effects are significant. This approach has been the subject of many investigations since its publication. Numerous elaborations and variations of the Jones' formula exist, each valid under different assumptions (e.g. Henderson, 1966; Di Silvio, 1969; Gergov, 1971; Birkhead and James, 1998; Fenton and Keller, 2001; Perumal et al., 2004; Petersen-Øverleir, 2006). However, these methods focus on unsteadiness effects solely, thereby disregarding variable backwater effects. To account for backwater effects, techniques involving two stage gauging stations at adjacent cross-sections were developed (Fenton and Keller, 2001; Arico et al., 2008; Dottori et al., 2009). However, the application of formulas using simultaneous water level measurements was criticised for lowland meandering rivers, where the water level gradient can be smaller than the measuring accuracy of the gauge (Koussis, 2010; Dottori and Todini, 2010). To overcome the practical concerns and limitations own to the twin level gauges approaches, Hidayat et al. (2011) proposed the use of velocity measurements in combination with only one stage measurement.

Aside from the more theoretical methods originating from the dynamic flow equations, soft computing techniques have recently been applied for the modelling of rating curves. A significant advantage of these approaches is that they do not impose a rigid model structure for transforming the input into an output. The performance of different types of artificial neural networks and the more complex adaptive neuro fuzzy inference systems were examined and

compared in multiple earlier studies that focus on rating curves (e.g. Deka and Chandramouli, 2003; Lohani et al., 2006). Such techniques were already discussed and used extensively in Chapters 2, 3 and 4. However, their "black box" nature and proneness to overfitting are two major disadvantages to be considered when applying these approaches, despite the satisfactory results they yielded in the reported case studies.

5.3 Data

5.3.1 Case Study 1: Marke River

Simulation results of a detailed full one-dimensional hydrodynamic model of the Belgian river Marke were employed in this study. The hydrodynamic model was built using the InfoWorks RS software and discussed already in §2.6.1. This InfoWorks RS model is also employed in Chapter 6. The river's flow is controlled by multiple gated weirs whose crest levels are regulated via fixed logical rules. Floodplains are also incorporated in the model based on field surveys and digital elevation data. Both these floodplains and the controllable hydraulic structures will greatly influence the rating curves as discussed below. Through the use of such full hydrodynamic model, measurement uncertainties are avoided, but effects of backwater, unsteadiness and variable storage can be examined.

Five events with hourly data of one month or less are investigated: 'Mar2008' (25/02 – 27/03/2008; 745 samples), 'Feb2010' (10/02 - 14/03/2010; 769 samples), 'Aug2010' (16/08 – 31/08/2010; 356 samples), 'Nov2010' (01/11 – 30/11/2010; 673 samples) and 'Jan2011' (20/12/2010 – 19/01/2011; 721 samples). During these periods, significant floods occurred along the river. All the stage and discharge data used in the examined approaches for the modelling of rating curves are retrieved from simulation results of the InfoWorks RS model at one location, at which flooding occurs if the stage surpasses 37.4 m. The latter two events show strong hysteretic behaviour, as illustrated in Figure 5.2. The results during the model's initialization phase (approximately the first 24 hours) were removed prior to this research. The events 'Mar2008' and 'Nov2010' were combined into one data set and used for calibration. The aim of selecting complete events as calibration data is to capture a wide range of conditions, including steady state situations and rapidly varying water levels that cause hysteretic effects. Since the examined approaches are data-driven, a runoff event resulting in a looped rating curve ('Nov2010') should be included in the

calibration data to ensure that the obtained models can emulate such behaviour. The other three act as validation data in all analyses. After the model construction, the performance is assessed and reported for both calibration events individually.

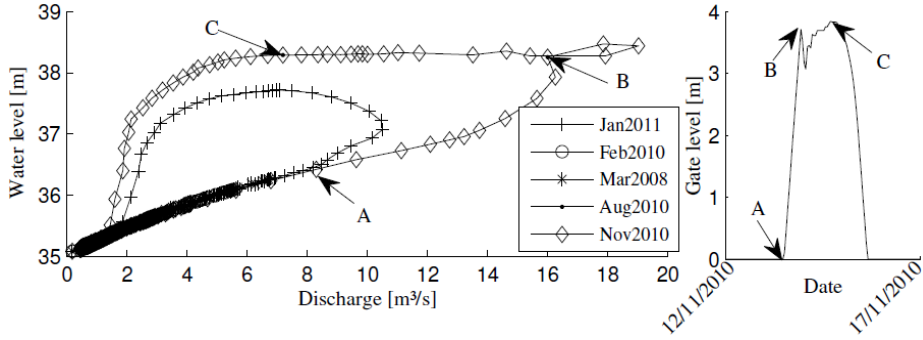


Figure 5.2: Discharge - stage couples for the 5 examined events (left) and the gate level of the weir downstream of the selected location with the gate's sill level as zero reference (right) for (part of) the event 'Nov2010'.

Approximately two kilometers downstream of the selected location where the stage and discharge data were acquired, a gated weir with changing gate level influences the flow. The gate opening depends via a complex blend of fixed rules on several observed water levels, including one upstream and one downstream of the controlled gated weir. In general, when the water level upstream (i.e., at the examined location) rises, the gate level augments as well, forcing the water to flow into the neighbouring floodplains. A safety measure is also included which lowers the gate level if the stage surpasses a specified threshold value. This varying gate level will affect the flow in the river significantly via backwater effects, and this is probably one of the main causes of the hysteresis shown on the left in Figure 5.2. The gate level between 12/11/2010 and 17/11/2010 is shown on the right in Figure 5.2. Before and after this period, the gate level is zero, which means the gate is fully opened. The rapidly fluctuating gate level between B and C (Figure 5.2) causes an additional complexity for the modelling of the rating curve. A similar course of the gate level occurred during the event 'Jan2011', whereas the gate remained always fully opened during the other three events. There are no gated weirs upstream of the selected location.

The accuracy of the flow predictions would likely increase if gate level readings are included in the prediction model. However, this research addresses a more general problem, namely that of mimicking flow data when only stage

measurements are available as is often the case in practice. Hence, information about what causes the backwater effects (in this case the varying gate level) is deliberately not incorporated in the rating curve models in this research. Therefore, the examined approaches should handle backwater effects caused by other factors, such as tributaries or ponds, in a similar way.

5.3.2 Case Study 2: Dender River

All the data used for the second case study are extracted from simulations with a MIKE11 model of the Dender River in Belgium. The calibration and validation of the hydrodynamic model was executed in the same manner as the InfoWorks RS model described for the first case study. This MIKE11 model is also used in Chapter 6 to configure a conceptual model, which was then used to quantify the impact of retention basins on river floods. The hourly simulation results of five different events of approximately one month are analyzed: 'Jan1995' (15/01 – 15/02/1995; 714 samples) and 'Nov2010'(1/11 – 30/11/2010; 714 samples), and synthetic runoff events with return periods of 25, 100 and 250 years (denoted as 'T25', 'T100' and 'T250' with 714, 714 and 840 samples respectively). These synthetic events represent input hydrographs for given return periods. See Willems et al. (2002) and Willems (2013) for more details on these synthetic events and the simulations. The historical flood 'Jan1995' and the synthetic event 'T100' were combined into a calibration set, the other three were used for the validation. By using an event outside the calibration range (i.e. 'T250') as validation data, the extrapolation capabilities of the methodologies will be assessed.

The location in the model where stage and discharge couples are collected is situated approximately 10 kilometers upstream of the confluence of the Dender with the Scheldt River. A combination of a gated weir and two vertical sluices inhibit the flow from the Scheldt River into the Dender during high tide in the Scheldt River. The hydraulic structures are reopened after the maximum stage of the high tide is reached. Hence, water coming from upstream in the Dender River is temporarily retained during high tide and released into the Scheldt River during low tide. Thereby, the stage time series in the downstream part of the Dender show a 12 hourly oscillation with the tide. Between the selected location and the group of hydraulic structures controlling the flow at the confluence with the Scheldt River, a second smaller gated weir has been installed along the Dender River. This weir prevents the stage of lowering below a specified threshold,

making navigation possible at all times. The time series of the stage and discharge for the event 'T100' is shown in Figure 5.3.

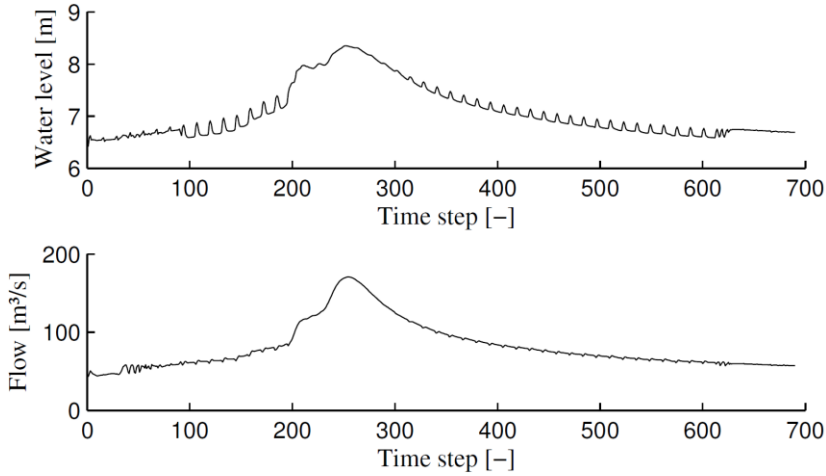


Figure 5.3: Time series of the water level (top) and discharge (down) for the event 'T100' at the selected location.

For stages between 6.5 and 8 metres, the influence of the tide can clearly be seen. When the stage in the Scheldt River surpasses that of the Dender River, the hydraulic structures at the confluence close and the water level quickly rises. When the stage in the Scheldt River lowers again, the structures are reopened and the water is released into the Scheldt. The associated drops of the water level cannot be seen in Figure 5.3, since a second gated weir rises to maintain a specified minimum stage. Because of backwater effects occurring between the selected location where the stage-discharge couples are extracted and the location that determines the gate level of the smaller gated weir, the selected threshold value for the minimal stage cannot be seen in Figure 5.3. The synergetic effect of the different hydraulic structure and the tide yields a rating curve with many small hystereses (see Figure 5.8).

5.4 Modelling approaches

5.4.1 Simple rating curve

The simple rating curve (SRC) approach uses a one-to-one mapping of river stage to discharge estimates. The curve disregards any effect that downstream backwater and unsteadiness might have, as well as all other influencing factors

described earlier. Commonly, the following power-law function is used in hydrometric practice (Herschy, 1978; Dymond and Christian, 1982; ISO 1998):

$$Q(t) = a(h(t) - b)^c \tag{5.1}$$

with Q the discharge [m³/s], h the stage measurement [m], a and c constants [-] and b a datum correction [m]. In this study, the three unknown parameters are optimized in a least squares sense via a trust-region-reflective algorithm, which is based on the interior-reflective Newton method (Coleman and Li, 1994, 1996), using the non-linear equation above. Note that experiments showed that performing a (natural) logarithmic transformation followed by a linear least squares optimization delivered inferior results compared to the (more complex) non-linear optimization. This is probably due to the errors which are not log-normally distributed.

5.4.2 Rating curve with state dependent parameter estimation

Jones' formula (1916) is a well-known and established method for calculating the discharge from stage measurements at a single location, which also incorporates effects of unsteadiness:

$$\frac{Q(t)}{Q_r(t)} = \sqrt{1 + \frac{1}{c(t)\bar{S}}\frac{\partial h(t)}{\partial t}} \tag{5.2}$$

with $Q_r(t)$ the steady rated discharge [m³/s] corresponding to stage $h(t)$, $\bar{S}$ the mean slope of the stream [-] and $c(t)$ the kinematic wave speed [m/s]. The steady rated discharge can be approximated via Eq. 5.1 using stage measurements and knowledge of the cross-section's geometry (e.g. Fenton and Keller, 2001). The expression is accurate for waterways with flood waves moving as kinematic waves. Although this methodology can be successful under certain conditions (Perumal et al, 2004.), the kinematic wave equation fails to capture backwater effects (e.g. Tsai, 2005; and see §1.2.2.1).

This research investigates a similar approach as that of Jones by multiplying the outcome of an SRC (cf. Eq. 5.1) with a time-varying (TV) state dependent (SDP) parameter. Hence, parameter a in Eq. 5.1 is made time-dependent to account for the dynamic effects encountered, while parameters b and c remain time-invariant. The latter two parameters are determined as in the SRC approach, before the

estimation of the TVPs. Instead of the rigid equations proposed by Jones and others (see above), the TVPs allow for a more flexible model.

Because the time variability is not known prior to the analysis, a stochastic model is associated with each TVP during estimation. Jakeman and Young (1979, 1984) assume a generic Generalized Random Walk process to describe the stochastic evolution of each parameter. This association simply acts as a statistical device to allow for the estimation of parametric change. Next, a Kalman Filter is employed in a forward pass (i.e., the data is handled sequentially in temporal order) for parameter estimation, followed by a backward pass through an optimal recursive smoothing (see e.g. Bryson and Ho, 1969; Norton, 1986). The latter is done via a Fixed Interval Smoothing (FIS) algorithm. This way, the obtained estimate at sample *t* is based on all available data instead of only the data up to sample *t*. Hence, the phase lag associated with the forward-pass filtering estimate is removed. In this study however, it is likely that the TVP itself depends in turn on the input variables (see for instance Eq. 5.2), yielding a truly non-linear system with rapid variations. In such case, the TVP estimation technique described above has to be altered to include the inherent state dependencies. Prior to FIS estimation, the data series are sorted in ascending order to obtain a smoother course, eliminating partly the quick temporal changes. If multiple SDPs are involved, each parameter is sorted (and estimated) in turn applying an iterative backfitting procedure. This results in a non-parametric relationship between the rapid parameter variations and the associated state. Finally, these relationships can be parameterized via various techniques such as simple regression, artificial neural networks and adaptive neuro fuzzy inference systems. In this study, several linear relationships are used for this purpose to ensure transparency. The above discussed procedure is described entirely in great detail in Young et al. (2001) and references therein. The mentioned algorithms are incorporated in the CAPTAIN Toolbox (Taylor et al., 2007; Young et al., 2007) for MATLAB.

As in Jones' method, the SDP correction is made dependent on the stage measurement and the difference in stages between two measurements:

$$Q(t) = \left[a_1\big(h(t)\big) + a_2\big(h(t) - h(t-1)\big)\right] \cdot (h(t) - b)^c \qquad (5.3)$$

with $a_1(h(t))$ and $a_2(h(t)-h(t-1))$ both SDPs that depend on the stage and difference in stages respectively.

This procedure fits in the data-based mechanistic philosophy (e.g. Young, 2003) in which model structure and parameters are chosen so that the model adequately represents the data, while also allowing for physical interpretation of the underlying system. The followed methodology should yield a parsimonious, efficiently parameterized model.

5.4.3 Artificial neural networks

Artificial Neural Networks (ANN) are employed increasingly for prediction in water management, including rainfall-runoff modelling (e.g. Shamseldin, 1997; Wu and Chau, 2011), river flow forecasting (e.g. Akiner and Akkoyunlu, 2012; Kalteh, 2013), and the modelling of rating curves (e.g. Ajmera and Goyal, 2012; Bhattacharya and Solomatine, 2005; Jain and Chalisgaonkar, 2000; Tawfik et al., 1997). ANNs are also incorporated in the developed conceptual modelling methodologies for rivers (see §2.3.1.2) and sewer systems (see §3.3.3.4) thanks to their flexibility and potential of emulation complex dynamics. This study uses ANN for the modelling of looped rating curves, so their potential and accuracy can be compared to the other evaluated methodologies.

This study employs the feedforward Multi Layer Perceptron (MLP) ANN, which is the most widely used type in hydrological studies (Maier and Dandy, 2000; Maier et al., 2010). Such an ANN consists of a number of neurons that map the (multidimensional) input to a crisp output via a transfer function, grouped in multiple layers. Each neuron in a layer is connected to all neurons in the following layer. The transfer functions of the neurons in the input layer are identity functions (Fausett, 1994), i.e. the output is equal to the input. Neurons in the hidden and output layers function in two stages. Firstly, the net input of each neuron is obtained by taking the sum of all inputs multiplied with their corresponding weights plus a constant weight (bias b). Next, the net input is transformed into the output via a non-linear function. Many types of transfer functions exist (see e.g. Fausett, 1994), but in this study the commonly applied logistic function is used.

The development of ANN models necessitates an increased attention to good practice, as they are built using available data and not explicitly based on underlying physical principles (Maier et al., 2010). Therefore, a brief description is given on the development procedure of the ANNs employed in this study. The ANN's inputs are the stage and the gradient of stage, as in the examined SDP-RC approach. The inputs are rescaled to the interval [0, 1] to ensure that all variables

receive equal attention during the training process, and to increase the efficiency of the training algorithms (Dawson and Wilby, 2001). Next, the calibration data is divided into training and validation subsets in a randomly unsupervised approach for 5-fold cross-validation. Hence, these training and validation subsets are used in the model optimization process. The Levenberg-Marquardt backpropagation algorithm is applied for network training with 100 epochs. One of the major concerns in ANN development is the selection of a suitable number of hidden layers and neurons. A balance should be found between the network complexity and the model's generalization ability (Maier et al., 2010). Networks with only one hidden layer are exploited in this research for the sake of simplicity and due to the successful results in other studies with these network architectures (see e.g. Chapters 2, 3, 6 and 8). The optimal number of neurons in the hidden layer was based on a trial-and-error procedure. Too few neurons restrict the model's capabilities, whereas too many increase the risk of overfitting. Finally, the performance of the obtained ANNs is tested on an independent set of data, denoted as 'testing' data (i.e. events 'Feb2010', 'Aug2010' and 'Jan2011' for case study 1, and 'Nov2010', 'T25' and 'T250' for case study 2). This research applies the algorithms from the Neural Network Toolbox for MATLAB.

5.4.4 M5' model tree

M5' trees were already introduced in Section 2.3.3, but are elaborated in this section in more detail. Decision trees are commonly applied in machine learning and data mining as a comprehensible knowledge representation. In general, a decision tree is a tree in which each branch node represents a choice between a number of alternatives and each leaf node represents a classification or decision (Hand et al., 2001). Regression trees emulate a non-linear function by a piecewise constant one, rather than dividing the data set in discrete categories. These trees are based on the well-known CART analysis (Breiman et al., 1984). Quinlan (1992) pioneered techniques for dealing with continuous-class learning problems by introducing "model trees" and the M5 ("M5" stands for "Model trees, version 5") learning algorithm. They have a conventional decision tree structure, but the leaves consist of linear functions instead of discrete class labels. Wang and Witten (1997) improved this technique and called it M5'. The latter technique will be applied in this research using the MATLAB toolbox developed by Jekabsons (2010).

In the first stage of model construction, a decision-tree induction algorithm is used to build the tree. The M5' algorithm splits the data based on the values of

predictive attributes, which minimize intra-subset variability. An in-depth description of the splitting criterion can be found in Wang and Witten (1997). After the tree has been grown, M5' computes a linear model for each interior node using standard regression. The attributes will be dropped one by one if they decrease the error estimate. Once a linear model is in place for each interior node, the tree is pruned back from the leaves. The modeller has to pick the minimum number of training data cases one node may represent. Too low values will reduce model's flexibility, whereas too large values increase the risk of overfitting. Hence, this parameter has a similar effect as the number of neurons in a hidden layer in ANN.

During model prediction, a smoothing procedure can be applied to compensate for the discontinuities between adjacent linear models. This process uses the leaf model to compute the predicted value, and then filters that value along the path back to the root, smoothing it at each node by combining it with the value predicted by the linear model for that node. The procedure is described in detail in Quinlan (1992).

Model trees have several advantages over alternative data mining techniques. The models can be represented by multiple if-then statements in combination with linear equations, making these models simple to understand and interpret. There is no implicit assumption of the underlying relationships between the predictor variables and the dependent variable. Hence, model trees are very flexible and can handle non-linear mapping. Moreover, little a priori knowledge or data preparation is required for the construction of the trees. Despite these clear advantages and the straightforward usage of the M5' algorithm, its use in water resources management is rather limited. Model trees have been applied in rainfall-runoff modelling (Solomatine and Dulal, 2003), flood forecasting (Solomatine and Yunpeng, 2004; Singh et al., 2010), statistical downscaling (Goyal and Ojha, 2012; Goyal et al., 2012) and also the modelling of rating curves (Ajmera and Goyal, 2012; Bhattacharya and Solomatine, 2005).

5.5 Results

5.5.1 Performance evaluation

A proper evaluation of the obtained models is vital to establish confidence in their performance. The position paper of Bennett et al. (2013) provides a

comprehensive overview of performance evaluation techniques and suggests a five-step procedure. Different methods are used to characterize the accuracy and suitability of the models. First, two (relative) residual methods are employed to assess model performance, namely the mean absolute percentage error (MAPE, Eq. 5.4) and the root mean square percentage error (RMSPE, Eq. 5.5). The use of percentage errors facilitates the comparison of different data sets since they are scale-independent. Both are expressed in percentages and have an ideal value of zero.

$$MAPE = \frac{100}{N} \cdot \sum_{i=1}^{N} \left| \frac{Q_{i,observed} - Q_{i,predicted}}{Q_{i,observed}} \right| \tag{5.4}$$

$$RMSPE = 100 \cdot \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(\frac{Q_{i,observed} - Q_{i,predicted}}{Q_{i,observed}} \right)^2} \tag{5.5}$$

with N the number of data samples and Q_i the river flow of time step i. The simulation results of the hydrodynamic models act as 'observed' discharges, whereas the output of each obtained model is used as 'predicted' discharge. The adjusted R^2 (denoted as $\bar{R}^2$, Eq. 5.7) value is also calculated, which is defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^{N} (Q_{i,observed} - Q_{i,predicted})^2}{\sum_{i=1}^{n} (Q_{i,observed} - mean(Q_{observed}))^2} \tag{5.6}$$

$$\bar{R}^2 = 1 - (1 - R^2) \cdot \frac{N - 1}{N - p - 1} \tag{5.7}$$

with p the number of parameters included in the model structure. The adjusted R^2 tries to tackle the phenomenon of generating higher accuracy metrics by including additional variables that have no explanatory power. Note that the adjusted R^2 value is always lower than the R^2 measure and has an ideal value of one. In addition, the observed and predicted values are visually compared by showing scatter plots together with a linear fit and R^2 values, or by concurrently plotting discharge-stage couples. The latter figures preserve the time-indexing, which is important when characterizing hysteretic behaviour.

Moreover, different model identification techniques exist that try to indicate whether a model has been overfitted by including too many parameters (e.g.

Jakeman and Hornberger, 1993). The Akaike Information Criterion (Akaike, 1974) for instance is a parameter-based metric that is often used to compare different models with each other, but is unable to evaluate the model in absolute terms. This criterion is generally very useful during the model identification phase, but is not examined in this study since the number of data points (observations) is several orders of magnitude larger than the number of parameters. This will inevitably lead to the full acceptance of the model with lowest residual sums of squares, and hence rejection of all other models, thereby not providing further insight into the accuracy and explanatory value of the different model structures. Secondly, the SDP-RC, ANN and M5' models do not have a prescribed model structure, making the number of parameters used highly case specific. Instead, a qualitative model evaluation is performed that focuses amongst others on the model adaptability and ease of the model build-up, and assesses the reliability and performance under different circumstances (see §5.6) (Nguyen et al., 2007; Krueger et al., 2012). Such qualitative and perhaps subjective analysis allows for the consideration of different elements that are often impossible to quantify, but that can be crucial to set-up accurate and reliable models that meet the intended aim.

5.5.2 Simple rating curve

5.5.2.1 Case study 1

First, the constant parameters of the Simple Rating Curve (SRC) approach are determined. Optimization yields 4.28 for coefficient *a*, 35.08 m for the datum correction *b* and 0.74 for *c*. The statistical properties of the fitted discharge series can be found in Table 5.1.

Due to the form of Eq. 5.1, this rating curve cannot capture hysteretic behaviour. The univocal rating curve endeavours the generally optimal result, leading to a compromise solution. Figure 5.4 shows the approximation for three events. Two events are not shown, because they are characterized by univocal relationships and all models emulate these well. Obviously, the events with hystereses ('Nov2010' and 'Jan2011') have significantly poorer fits than the other three events (see Table 5.1), which lie approximately on a single curve. The model manages to emulate the validation events relatively well, with statistical properties that even surpass some of the calibration events. The better fit for the Jan2011 event in comparison with Nov2010 is caused by the smaller hysteresis (see Figure 5.2).

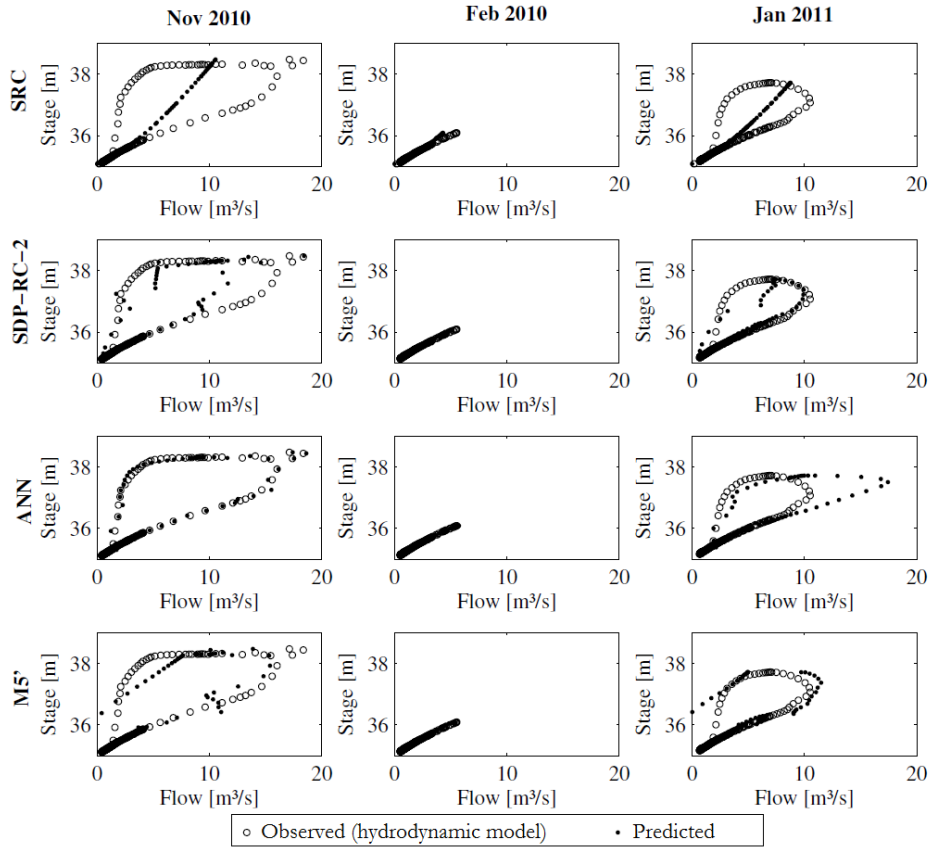


Figure 5.4: Fit of the different models in function of the observed stage for the calibration event 'Nov2010' (left) and validation events 'Feb2010' (middle) and 'Jan2011' (right) for the first case study (Marke River).

Table 5.1: Statistical properties of the Single Rating Curve (SRC), State Dependent Parameter – Rating Curve (SDP-RC), Artificial Neural Network (ANN) and M5' Model Tree (M5') for the first case study.

		Calibration Events		Validation Events		
		Mar2008	Nov2010	Feb2010	Aug2010	Jan2011
SRC	RMSPE (%)	12.01	29.79	10.53	16.59	19.38
	MAPE (%)	9.79	14.94	7.91	13.55	12.69
	adjusted R^2	0.94	0.74	0.96	0.97	0.82
SDP-RC-1	RMSPE (%)	5.66	14.86	4.13	14.01	13.04
	MAPE (%)	4.99	6.34	3.54	11.23	5.48

		Calibration Events		Validation Events		
		Mar2008	Nov2010	Feb2010	Aug2010	Jan2011
SDP-RC-2	adjusted R^2	0.99	0.87	0.99	0.99	0.92
	RMSPE (%)	7.28	12.38	4.96	15.97	10.89
	MAPE (%)	5.87	6.35	3.79	12.64	4.12
ANN	adjusted R^2	0.99	0.88	0.99	0.99	0.95
	RMSPE (%)	6.34	6.28	3.72	16.95	11.02
	MAPE (%)	4.99	4.20	3.06	13.04	4.38
M5'	adjusted R^2	0.99	0.98	0.99	0.99	0.77
	RMSPE (%)	4.11	12.42	3.97	7.97	7.57
	MAPE (%)	3.17	5.33	3.23	6.34	3.25
	adjusted R^2	0.99	0.92	0.99	0.99	0.96

5.5.2.2 Case study 2

Optimization of the SRC for the two calibration events yields 49.73, 5.51 m and 1.09 for parameters *a*, *b* and *c* respectively. The performance of the SRC is shown in Table 5.2. Again, the obtained SRC is a generally optimal solution. Because of the smaller hystereses in the data for case study 2 in comparison with case study 1, the adjusted R^2-values are closer to unity.

Table 5.2: Statistical properties of the Single Rating Curve (SRC), State Dependent Parameter-Rating Curve (SDP-RC), Artificial Neural Network (ANN) and M5' Model Tree (M5') for the second case study.

		Calibration Events		Validation Events		
		Jan1995	T100	Nov2010	T25	T250
SRC	RMSPE (%)	13.49	7.60	41.25	8.87	7.03
	MAPE (%)	9.47	6.14	21.70	7.11	5.14
	adjusted R^2	0.94	0.95	0.96	0.93	0.93
SDP-RC	RMSPE (%)	11.20	5.85	26.70	8.13	3.67
	MAPE (%)	8.18	4.51	13.69	6.77	2.62
	adjusted R^2	0.96	0.98	0.98	0.95	0.99
ANN	RMSPE (%)	10.03	5.58	43.51	8.46	4.86
	MAPE (%)	7.14	4.21	19.98	6.91	3.47

		Calibration Events		Validation Events		
		Jan1995	T100	Nov2010	T25	T250
M5'	adjusted R^2	0.96	0.98	0.97	0.95	0.96
	RMSPE (%)	9.76	4.56	15.49	8.20	3.86
	MAPE (%)	6.51	3.28	8.53	6.82	2.73
	adjusted R^2	0.95	0.98	0.98	0.93	0.98

The rating curves and the approximations obtained by the different models are not explicitly shown for this case study. Such figures would be unclear because of the high variability and scatter of the rating curves (see e.g. Figure 5.8). Instead, Figure 5.5 shows the observed (i.e. data retrieved from hydrodynamic simulations) versus the predicted discharges for the different methodologies. A least squares linear fit is indicated by the full line, together with the equations and R^2-values. It is clear that the SRC approach underestimates the higher discharges (e.g. event 'T250') significantly. The scatter regularly exceeds the ±10% bounds (indicated by the dashed lines).

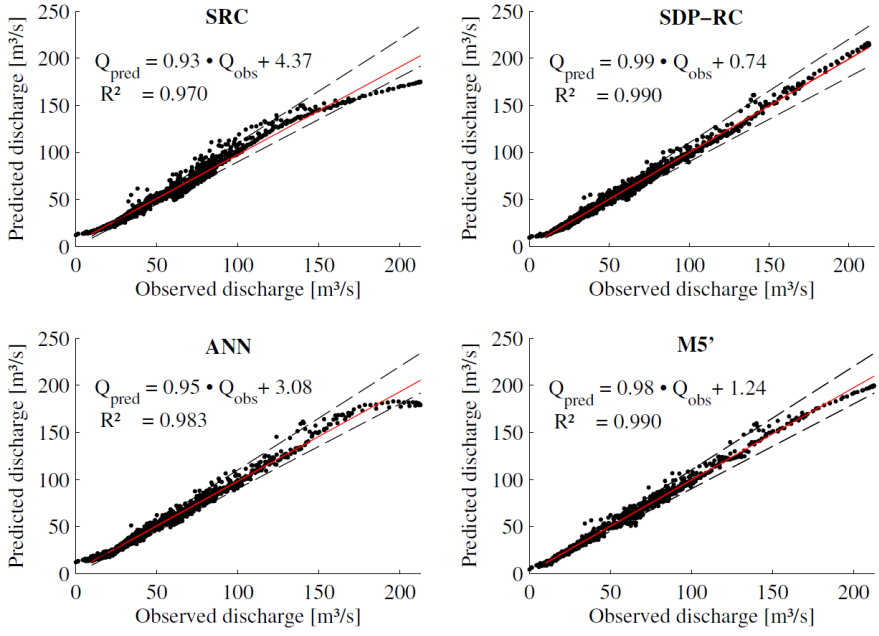


Figure 5.5: Observed (extracted from hydrodynamic simulations) versus the predicted discharges of the four examined approaches for all events of case study 2. The observed discharges ±10% are indicated by the dashed lines. The full lines indicate a linear fit, together with the equations and R^2 values.

5.5.3 State Dependent Parameter Rating Curve (SDP-RC)

5.5.3.1 Case study 1

Next, the in this research proposed State Dependent Parameter Rating Curve (SDP-RC) approach is examined. The parameters *a*, *b* and *c* of Eq. 5.3 are optimized in the same manner as in the SRC approach, but only the event of Mar2008 was employed as calibration data. The discharge-stage couples exhibit a univocal relationship for this smaller event (see Figure 5.2) and are therefore considered to approximate a steady rating curve. Thereby, the SDP correction is entirely aimed at tackling the hysteresis caused by variable factors. If an event with hysteresis was included in the calibration data, the SRC optimization would yield a curve that deviates from the steady rating curve. This would impede the physical interpretation of the SDP outcome. Optimization yields 34.97 m for the datum correction *b*, a value of 1.25 for exponent *c* and a constant value of 4.80 for *a*. This latter value is solely used for comparison with the sum of $a_1(h(t))$ and $a_2(h(t)-h(t-1))$ (cf. Eq. 5.3) obtained after the SDP analysis.

The following analysis is performed on the complete calibration data, namely events 'Mar2008' and 'Nov2010'. Integrated random walk processes, also known as smooth trend models, are associated with $a_1(h(t))$ and $a_2(h(t)-h(t-1))$ separately. The parameters of these integrated random walks are optimized in the first backfitting iteration and are kept constant for the remaining three iterations. The FIS algorithm delivers discrete values of the two time varying parameters shown in Figure 5.6, together with their 95% confidence intervals.

Several elements of the acquired SDP model corroborate its physical context. Parameter $a_1(h(t))$ shows an approximately constant value for stages between the minimum observed stage and 36.9 m. The average estimate of $a_1(h(t))$ in this interval is 3.06. The averaged value of $a_2(h(t)-h(t-1))$ for stage differences plus or minus 0.05 m is 1.68. Hence, these parts of both SDP are, as expected, very closely related to the alleged steady rating curve estimation. More striking is the steep drop of the $a_1(h(t))$ estimates for stages between 36.9 and 38.1 m. This corresponds almost perfectly with the augmenting gate level of the downstream gated weir. For stages surpassing approximately 36.6 m, the gate level rises (see point A in Figure 5.2). This varying gate level causes variable backwater effects, thereby inhibiting flow and creating hystereses. To mimic this hysteretic behaviour accurately, the second SDP $a_2(h(t)-h(t-1))$ is required. This SDP determines, as in the widely known rising/falling approach, on which branch the

data samples are located and applies a suitable correction factor. Finally, the fast rising part of $a_1(h(t))$ for stages above 38.1 m accounts for the more extreme values of the gate level. When the gate level reaches its maximum allowable position (3.85 m above the gate's sill level) and the stage keeps rising, it is likely that the inhibitory effect of the weir diminishes. At this stage, remember that numerous factors affect the river's state, including flood, the fluctuating gate level (see points B and C in Figure 5.2) and other unsteady dynamics. Despite all these challenging elements and uncertainties, the SDP analysis still manages to find estimates showing a trend.

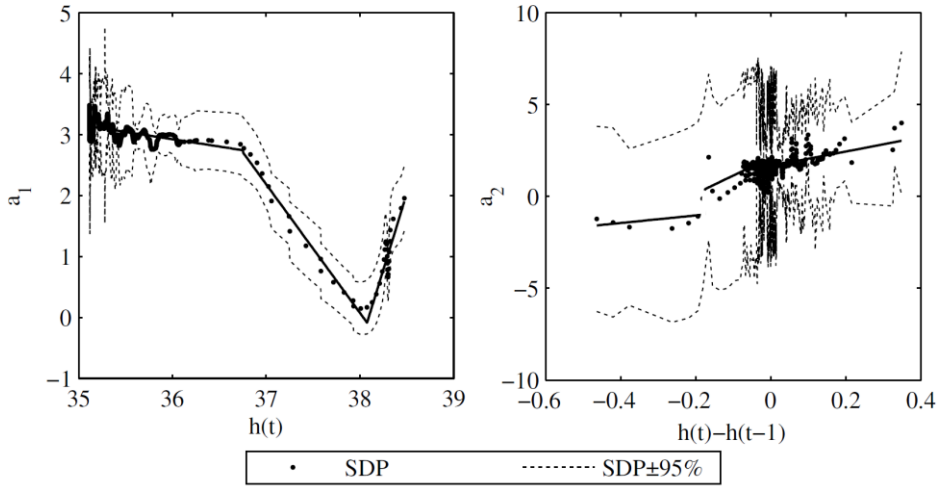


Figure 5.6: Parameters a1(h(t)) and a2(h(t)-h(t-1)) as identified using the SDP algorithm (dots) together with their 2x standard error confidence intervals (dotted lines) for the first case study. The parameterization was done using piecewise linear relationships for each SDP (solid lines).

The SDP estimates shown in Figure 5.6 yield an exceptional value of 0.998 for R^2 for the calibration data. Of course, these SDP estimates have to be parameterized based on their states before they can be exploited in a simulation model. For this purpose, simple piecewise linear relationships are used to accentuate the simplicity and preserve the transparency of the final model. After manually selecting breakpoints, linear relationships were fitted to the data between two adjacent breakpoints. The resulting piecewise linear relationships are shown in Figure 5.6. Due to the parameterization, the R^2 coefficient of the calibration data dropped to 0.896. Next, simulations were performed for all events using the obtained model. The results are summarized in Table 5.1 under 'SDP-RC-1'. It is

clear that the model emulates the full hydrodynamic results much better than the SRC approach. All adjusted R^2 values exceed 90%, except for the second calibration event.

Because of the rough parameterization of the SDP estimates, optimization of the parameters *b* and *c* in Eq. 5.3 might be beneficial given the values of $a_1(h(t))$ and $a_2(h(t)-h(t-1))$. Simultaneous optimization of parameters *b* and *c* and SDP estimation is not possible due to the implementation of the SDP algorithm. The non-linearities introduced by the time varying parameters can simply be imposed upon the model as time series, since their values can be calculated for the calibration data via the chosen piecewise linear equations. Optimization using the calibration data delivered values of 34.99 m and 1.19 for *b* and *c* respectively. The individual events were simulated again and the performance indicators are appended to Table 5.1 under 'SDP-RC-2'. Figure 5.4 shows the discharge-stage couples for the events 'Nov2010', 'Feb2010' and 'Jan2011'. The statistics of the two events with the poorest approximations ameliorated. Moreover, it is notable that the performance of the validation event showing hysteresis ('Jan2011') has better performance statistics than the calibration event with hysteresis, although its shape differs significantly from the supplied data during calibration. This indicates the good generalization capabilities and confirms the physical meaning of the model. In the falling branch of the rating curves, the predictions deviate from the observed values. This is caused by the higher uncertainty for the SDP estimates of $a_2(h(t)-h(t-1))$. Using solely a modest amount of observed discharge-stage couples, the SDP analysis managed to deliver an accurate model that can even account for the gate level variations downstream. The obtained model is highly transparent and has great generalization capability, which is proven by the good error statistics for the validation events. Note that it is likely that even better approximations can be acquired when more flexible methods are employed for SDP parameterization, such as neuro-fuzzy techniques, although the expected gain is limited. In such case, the structure of the non-linear system can still be used, but the SDP estimates are calculated using an adaptive neuro fuzzy inference system.

5.5.3.2 Case study 2

The same procedure as in case study 1 was followed for model calibration. Both calibration events, namely Jan1995 and High100 were used for the estimation of parameters *b* and *c* in Eq. 5.3, yielding of course the same parameters as in the SRC model (cf. §5.5.2.2). The SDP estimates are shown in Figure 5.7. The

averages of $a_1(h(t))$ and $a_2(h(t)-h(t-1))$ are also calculated and indicated on Figure 5.7 by the horizontal lines. Both averages are almost equal (24.771 and 24.768 for $a_1(h(t))$ and $a_2(h(t)-h(t-1))$ respectively) and their sum approximates the optimized value of a in the SRC approach (49.73, cf. §5.5.2.2).

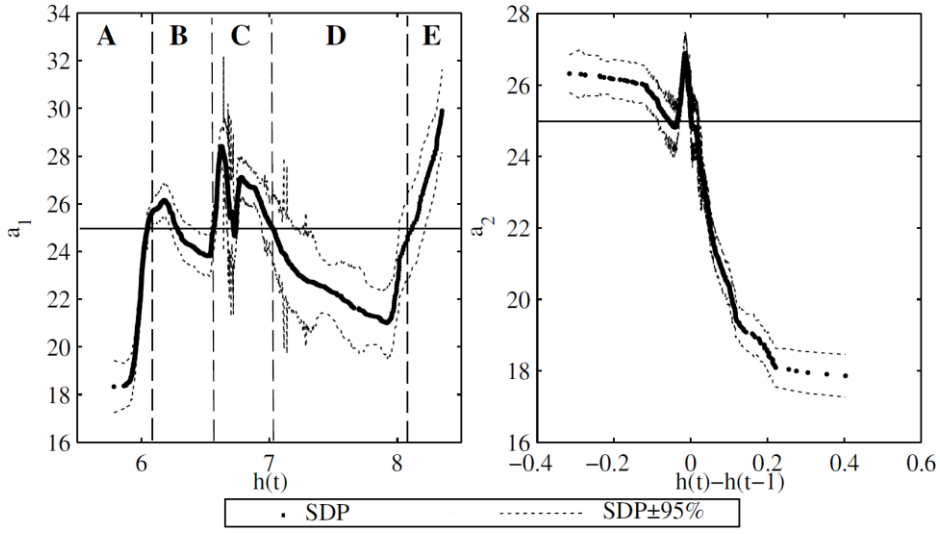


Figure 5.7: Parameters $a_1(h(t))$ and $a_2(h(t)-h(t-1))$ as identified using the SDP algorithm (dots) together with their 2x standard error confidence intervals (dotted lines) for the second case study. The averages of each parameter are marked by the horizontal full lines.

Again, a physical interpretation can be given for the estimates of each parameter. To facilitate the explanation, the estimates of $a_1(h(t))$ are divided into five zones, labelled 'A' to 'E' on Figure 5.7. As in the first case study, parameter $a_1(h(t))$ corrects the general shape of the rating curve. It tries to find a compromise solution through the hystereses as in the optimization of the SRC. In contrast to the SRC approach, parameter $a_1(h(t))$ depends on the stage and is therefore time-varying, allowing for a closer emulation of the data couples. The rating curve in the form of Eq. 5.1 with constant parameter a equal to the averaged sum of $a_1(h(t))$ and $a_2(h(t)-h(t-1))$ is shown in Figure 5.8 for the calibration data. It is clear that the rating curve systematically over- and underestimates the calibration data in some areas. In fact, an "S-shaped" (sigmoid) rating curve seems more appropriate, but the imposed model structure (see Eq. 5.1) impedes such form. This shape can be caused by the variable backwater and unsteadiness effects occurring for stages between roughly 6.5 and 8 m, or varying cross-section

geometries. In zone 'A', the constant rating curve overestimates the calibration data: most points are situated on the left of the rating curve. Hence, for stages lower than approximately 6.1 metres, the estimates of parameter $a_1(h(t))$ are below its average value. In zone 'B', no clear under- or overestimation can be noticed. Most data points of zone 'C' lie on the right side of the SRC approximation and therefore, $a_1(h(t))$ is slightly higher than the average value. Zone D is characterized by overestimations by the SRC model, which are corrected by the lower $a_1(h(t))$ estimates. For the highest stages of the calibration set, the backwater effects diminish, no hysteresis can be noticed and logically, the SRC approach underestimates the data couples significantly. The latter is corrected by the increasing $a_1(h(t))$ estimates.

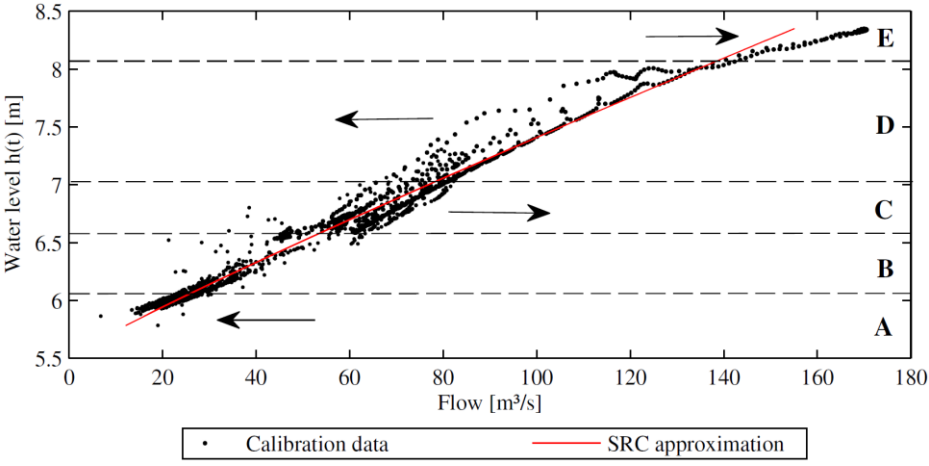


Figure 5.8: Calibration data ('Jan1995' and 'T100') together with the SRC approximation (cf. Eq. 5.1) with parameter *a* equal to the averaged sum of the estimates of $a_1(h(t))$ and $a_2(h(t)\text{-}h(t\text{-}1))$ from the SDP-RC analysis. The arrows indicate the effect of parameter $a_1(h(t))$ for each zone.

Again, as in the first case study, parameter $a_2(h(t)\text{-}h(t\text{-}1))$ accounts for the hystereses. A steep drop in the $a_2(h(t)\text{-}h(t\text{-}1))$ estimates can be observed for rapidly increasing stages with narrow uncertainty bounds. This corresponds with the earlier explained sudden closure of the hydraulic structures when the stage in the Scheldt River surpasses that of the Dender due to the tide. When the hydraulic structures prevent water from flowing into the Scheldt, the stage rises fast and backwater effects occur. Hence, the higher $a_2(h(t)\text{-}h(t\text{-}1))$ estimates correct for this phenomenon. It can be argued that such sudden rises of the stage could also occur during very extreme flood waves. In such case, the given explanation does

not apply and a different behaviour could be noticed. Therefore, the performance of the SDP-RC approach will be validated on an extreme synthetic runoff event with a return period of 250 years ('T250') as well. For faster declines of the river stage, the situation is a bit fuzzier. In general, estimates of $a_2(h(t)-h(t-1))$ lie above the average value, which corresponds to the decreasing backwater effects when the hydraulic structures are reopened. However, the estimates fluctuate significantly for differences in subsequent stage recordings just below zero. This likely originates from the complex interactions between the gated weir in the vicinity that prevents the stage of dropping below a specified threshold on the one hand, and the hydraulic structures at the confluence downstream on the other hand.

The SDP estimates result in a value for R^2 of 0.986. As in the first case study, the parameterization of the estimates was done using a piecewise linear relationship. A new segment of this relationship for $a_1(h(t))$ starts at the beginning of each zone shown in Figure 5.7 (the breakpoint between zones 'D' and 'E' was shifted slightly to the left). The parameterization of $a_2(h(t)-h(t-1))$ includes four linear equations (a first one up to -0.05m, a second and third to emulate the peak around zero and the last one starting at 0.15m.). Next, parameters b and c (cf. Eq. 5.3) were optimized again using the calibration data to account for errors introduced by the parameterization of the estimates. The performance criteria for the two calibration and the three validation events are appended to Table 5.2. Notice the good error statistics for the largest event ('T250'), indicating the great extrapolation capabilities of the model. This is also confirmed by the comparison shown in Figure 5.5. The steep drop of $a_2(h(t)-h(t-1))$ for fast rising stages does not affect the performance negatively. The least squares fit shown in Figure 5.5 has a slope that is very close to unity.

5.5.4 Artificial neural networks

5.5.4.1 Case study 1

Calibration of the ANN was done using 5-fold cross validation. The number of hidden neurons was varied between 2 and 7 and training was conducted with 100 epochs. The results showed that neural networks with 6 hidden neurons led to the lowest sum of the RMSEs of the 'training' and 'validation' subsets (both part of the calibration data). The ANN contains 25 weights and biases, which is much lower than the number of samples used for training (1155). Next, the

performance was examined on the independent 'test' set (i.e. events 'Feb2010', 'Aug2010' and 'Jan2011').

The ANN approach delivers outstanding predictions for the calibration events, with values for adjusted R^2 close to unity (see Table 5.1). However, the calibration event of Jan2011 has poor predictions, with even larger errors than the SRC yielded. This indicates that the model could have been overfit, despite the carefully conducted calibration procedure. The ANN's very flexible architecture makes this type of model prone to overfitting. More (diverse) data during training and accompanying cross-validation is absolutely required to obtain models with better generalization abilities.

5.5.4.2 Case study 2

The same procedure was followed as in case study 1. The cross-validation led to the inclusion of 6 hidden nodes. The error statistics are included in Table 5.2. Again, the ANN model emulates the calibration data very well, although the differences with the SRC and SDP-RC approaches are smaller in comparison to the first case study. The error statistics of the independent 'test' sets (i.e. events 'Nov2010', 'T25' and 'T250') are close to each other and adequate, but are all surpassed by those of the SDP-RC approach. The emulation of the event outside the calibration range is poor as indicated by Figure 5.5: the highest discharges are significantly underestimated by the ANN approach.

5.5.5 M5' model tree

5.5.5.1 Case study 1

Finally, the M5' model tree approach was tested. As in the latter two approaches, the stage and gradient of stage (*h(t)-h(t-1)*) are used as inputs. The minimum number of training data cases one node may represent was set to 5. A smoothing coefficient of 12 was applied as proposed by Wang and Witten (1997). Training resulted in a model tree with 10 leaves (consisting of 27 parameters in total), each representing a linear model. The performance criteria are listed in Table 5.1. This model yields in general the best results of all examined approaches. As in the SDP-RC approach, the error statistics of the validation event with hysteresis ('Jan2011') surpass those of the calibration data.

Close inspection of the model's architecture and parameters learns that most data samples are situated on a single curve, denoted earlier as the steady rating curve.

It is therefore logic that the M5' algorithm employs several linear relationships to approximate this curve. Small errors on individual estimates in this region are quickly enlarged in the final error statistics because of the great number of samples. The second input, namely the difference in subsequent stages, is only of importance for stages surpassing 36.32 m. For these stages, the second input argument determines, as in the SDP-RC, whether the data sample is located on the rising or falling branch of the hysteresis. Because of the inclusion of the second input itself in the linear relationships, the shape of the hysteresis can vary. A disadvantage of this approach is that 'discontinuities' can occur in model prediction when the model switches from one relationship to another. The smoothing factor should tackle this problem, but the jumps can still be noticed (e.g. the rising branch of 'Jan2011' in Figure 5.4). Secondly, the falling branch is approximated too roughly by the M5' algorithm. The user could however manipulate the relationships manually, although obtaining an improved model in this way is not a sinecure. Despite the mentioned shortcomings, the automation and objectivity of the model construction and the greater flexibility puts this approach well in favour of the ubiquitously used rising/falling approach.

5.5.5.2 Case study 2

The error statistics of the M5' approach are added to Table 5.2. Training yielded a model with 92 leaves (i.e. linear equations) that consists of 199 parameters. The satisfactory performance of the trained M5' model for the validation events indicate that the trained M5' models are probably not overfitted. The highest discharges of the validation events are slightly underestimated as shown in Figure 5.5, but the overall agreement is very good as indicated by the high R^2-value. Because of the high number of equations, the transparency is greatly reduced, which was one of the major advantages of this kind of model. The number of leaves can be lessened by increasing the minimum number of data cases one node may represent.

5.6 Performance assessment using measurement data

The presented research obviates several difficulties by relying on simulation results of detailed models. This discussion addresses how the examined methodologies will behave in situations that are often encountered in practice, namely when only a small number of samples is available, the models are used for extrapolation, and the data set contains discontinuities or outliers.

If merely a limited data set is disposable, purely data-driven models, such as ANN and M5' model trees, are affected most. These approaches solely depend on data to identify the model structure and calibrate the parameters instead of relying on physical principles. If too little information is presented during model built up, these models are easily overfitted. This is demonstrated by the ANN construction in the first case study, where the performance of a validation event ('Jan2011', see Table 5.1) is even poorer than that of the SRC approach. The SDP-RC approach is less affected by this issue. The obtained SDP estimates are also data-driven, but can easily be interpreted. The modeller can still incorporate expert knowledge and alter the proposed relationships.

In the second case study, the extrapolation capability of the different approaches was tested by validating to an event which is much more extreme than the calibration set ('T250'). It is shown that all models manage to yield adequate predictions. As with all regression and data-driven methods, a prerequisite is the inclusion of the most important principles affecting the relationships in the calibration set. Thanks to its rather simple model structure, the behaviour of the SDP-RC models is easier interpretable and predictable in extreme conditions. Note that because of the limited number of inputs, simple two-dimensional grids or tables can be created that show the outcome of each model for a given input set.

Discontinuities in the calibration set do not pose much trouble. In contrast to auto-regressive models, a precise estimate of the discharge at the first time step is not required. Only the difference in subsequent stage measurements, which is used as an input, should be corrected. If the number of discontinuities is limited with respect to the amount of calibration data, setting this input to zero will likely have negligible effects on model performance.

To test the influence of outliers, 5% of the discharges of case study 2 was increased or decreased by 40%. Figure 5.9 shows the adapted calibration set used to test the four methodologies. The same model built up procedure was followed as for the two case studies (see §5.5). The performance of the retained models is tested on the original (unaltered) data set. The resulting adjusted R^2-values are shown in Table 5.3, together with their change compared to the corresponding models that were calibrated using the original data set (see Table 5.2).

Table 5.3: Adjusted R^2-values of the different models (trained on a calibration data set with artificial outliers) for the original data series. The differences in adjusted R^2 values in comparison with the models that were calibrated using the original data set are indicated between brackets ($\cdot 10^{-3}$).

	Calibration Events		Validation Events		
	Jan1995	T100	Nov2010	T25	T250
SRC	0.94 (-0.7)	0.95 (+0.2)	0.96 (-1.5)	0.93 (-0.2)	0.93 (+3.1)
SDP-RC	0.95 (-1.9)	0.98 (-1.1)	0.98 (+0.2)	0.95 (+0.7)	0.99 (+1.3)
ANN	0.96 (-4.0)	0.98 (-5.0)	0.97 (+1.8)	0.95 (-3.1)	0.98 (+19.7)
M5'	0.95 (+0.6)	0.98 (-1.1)	0.97 (-10.7)	0.93 (-0.8)	0.97 (-6.4)

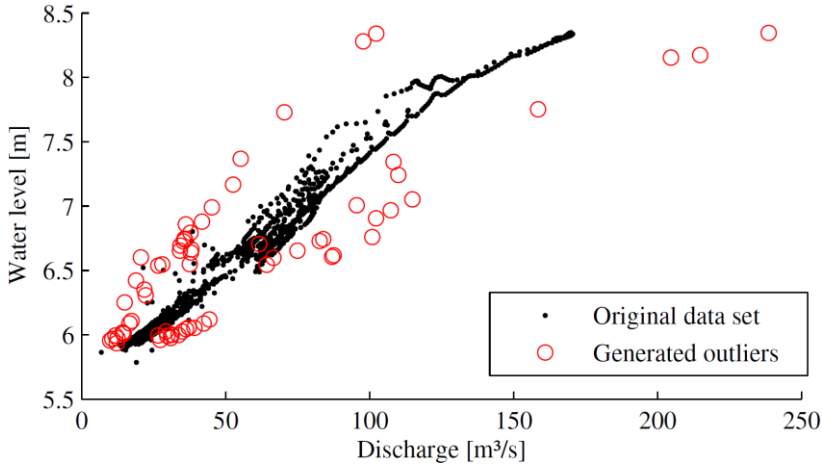


Figure 5.9: Part of the original calibration set (dots) together with the generated outliers (circles).

As expected, the performance of the SRC model is almost equal. In the SDP-RC approach, the outliers only have a paltry effect on the estimates of parameters $a_1(h(t))$ and $a_2(h(t)\text{-}h(t\text{-}1))$, but they enlarge the uncertainty bounds. As in case study 2, five and four linear relationships are used to parameterize respectively $a_1(h(t))$ and $a_2(h(t)\text{-}h(t\text{-}1))$. Because almost the same estimates were obtained, the performance hardly changes. Cross-validation led to the inclusion of only four hidden nodes in the ANN instead of 6 (see §5.5.4.2). Therefore, the ANN is less tailored to the input data and manages to emulate the validation events much better. The trained M5' tree comprises 79 leaves. The calibration events were mimicked well, but the performance of the validation events deteriorated. Hence,

the overall effect of the outliers is small, but they affect the ANN and M5' models more.

5.7 Discussion and conclusions

Rating curves are often characterized by looped trajectories due to various influences. Neglecting this hysteretic behaviour can result in large errors in flow predictions and thereby affect all applications that rely on these predictions. Conventional modelling techniques cannot emulate hystereses, but various theoretically well-founded approaches exist that account for such complex behaviour. These techniques incorporate additional measurements, such as stage readings at an adjacent location or velocity estimates, but this additional information is often unavailable. Therefore, this study aims to emulate variable rating curves when only discharge and stage measurements at one location are present. The performance of physically based models on the one hand, namely the simple rating curve (SRC) approach and a proposed expansion with state dependent parameters (SDP-RC), and expert systems on the other hand, specifically artificial neural networks (ANN) and the rarely used M5' model trees, are compared quantitatively and qualitatively. Data used in the experiments are gathered at two locations in detailed hydrodynamic models with rating curves that are strongly affected by hystereses.

The conducted experiments indicate that the commonly used SRC approach delivers inferior results in comparison with the other tested methodologies. As expected, the SRC models are not capable of emulating the hysteretic behaviour. In contrast, a proposed time-varying form of the conventional rating curve with state dependent parameters (SDP-RC) manages to represent complex rating curves with hysteretic behaviour. The SDP-RC models deliver accurate results (see Table 5.1, Table 5.2, Figure 5.4 and Figure 5.5) while remaining highly transparent. The presented methodology yielded physically interpretable models, as demonstrated by analyzing the SDP estimates in detail. The often applied ANNs outperform the SDP-RC models for the calibration data, but deliver poorer predictions for the validation data despite a carefully followed calibration procedure. The complex black box model structure and lack of physical background make this approach sensitive to overfitting. When only a limited amount of data is present and the exerted behaviour between the states is complex, creating an ANN with good generalization capabilities can be very difficult. The M5' model trees produced in general slightly better predictions than

the SDP-RC approach and have a fairly comprehensible model structure, which consists of a simple set of linear equations at the leaves. However, the additional gain with respect to the SDP-RC models does not justify the lower transparency and additional complexity. Therefore, the SDP-RC is selected here as the preferred method among the tested models for modelling rating curves affected by hysteresis when only (a modest amount of) stage and discharge measurements are available.

The research is entirely based on simulation results from detailed hydrodynamic models that solve the full de Saint-Venant equations. This way, boundaries can be picked and measurement uncertainties are avoided. Special attention was paid to practical concerns, such as the presence of discontinuities and outliers in the calibration time series, and the extrapolation capabilities of the models. Clearly further testing on other sites is necessary and we therefore cannot claim that the SDP-RC approach will be appropriate in all cases. Also, additional ANN structures (e.g. radial basis functions) and more complex expert systems (e.g. adaptive neuro fuzzy inference systems; see also Chapter 4) could be considered. Some important elements affecting rating curves are not investigated in this research, such as erosion and sedimentation of the channel, changes in roughness and the influence of vegetation. Future research could focus on these aspects.

The investigated model structures were also incorporated in the general conceptual modelling approaches presented in Chapters 2 and 3. The simple rating curve (SRC) approach of this study is the default method to model rating curves (see also §2.3.3.1). The M5' model trees can be used for the modelling of rating curves and hypsometric curves (§2.3.3.1), while ANNs can be employed for modelling flow (§2.3.1.2 and §3.3.3.4) and rotating water surface profiles (§2.3.3.2).

6

Application 1: Impact quantification of flood control basins in the Dender catchment

6.1 Introduction

Floods cause worldwide huge economical losses and it is likely that floods will even strike increasingly and more destructively due to urbanization and climate change (e.g. IPCC 2012; Willems et al., 2012; Arnbjerg-Nielsen et al., 2013). As discussed extensively in the introductory chapter, models with a limited calculation time will play a crucial role in future river and flood management (see §1.1-§1.3).

This chapter uses conceptual river models for a case study in flood management to illustrate the use and several advantages of conceptual models that are configured using the newly developed approach. In the investigated case study, the reduction of water levels and discharges is quantified due to the installation of additional and the enlargement of existing retention basins along the Marke River in Belgium. In addition, different gate operation controls are explored. It is essential to assess the structural adaptations on the basis of a long term simulation, followed by a statistical analysis of the results, since this allows

accounting for antecedent conditions. This approach also allows calculating the return period of the response of the system while avoiding statistical pre-processing of the inputs. However, performing long term simulations requires computationally efficient models that can simulate a long time series in a limited amount of time. If management strategies require optimization, it is likely that such long term simulations must even be performed a number of times. Therefore, a conceptual model is used that was configured using the conceptual modelling approach for rivers presented in Chapter 2.

This chapter is outlined as follows. First, the case study area, selected retention basins and different gate control strategies are discussed. Section 6.3 presents the calibration and validation results of the derived conceptual models. These models represent the current state of the rivers Marke and Dender. After implementation of the retention basins in the conceptual models, a partial recalibration was performed. Such recalibration is necessary since the conceptual models are based on empirical relationships and therefore cannot predict the flow in the altered system. The derived conceptual models of the Marke and Dender were coupled and used to simulate the historical flood event of November 2010 and a 36-years rainfall time series. These results are discussed in Section 6.4 together with an interpretation of the obtained results. By statistically analyzing the simulation results of the models representing the current and future states of the rivers, it was possible to quantify the reduction of flood frequency and magnitude.

This chapter is based on the following publication:

> WOLFS, V., WILLEMS, P. (2015). QUANTIFICATION OF IMPACT OF RETENTION BASINS ON RIVER FLOODS IN THE DENDER CATCHMENT IN BELGIUM USING COMPUTATIONALLY EFFICIENT MODELS. 36TH IAHR WORLD CONGRESS, DELFT/THE HAGUE, THE NETHERLANDS, 29 JUNE – 3 JULY 2015.

6.2 Case study

6.2.1 Study area

The Marke is a non-navigable river in Belgium that mouths in the Dender River (see Figure 6.1). Both rivers are regularly affected by floods. The discharge of the Marke and Dender Rivers amount to respectively 2 m^3/s and 5 m^3/s at their confluence under normal flow conditions, which can rise to approximately 20

m^3/s and 100 m^3/s during periods of severe flooding. The flow of the Marke and Dender Rivers is controlled by several movable weirs. The movement of the gated weirs along the Marke River is determined by a set of logic controls that rely on multiple water level measurements, while those of the Dender River are controlled manually based on experience.

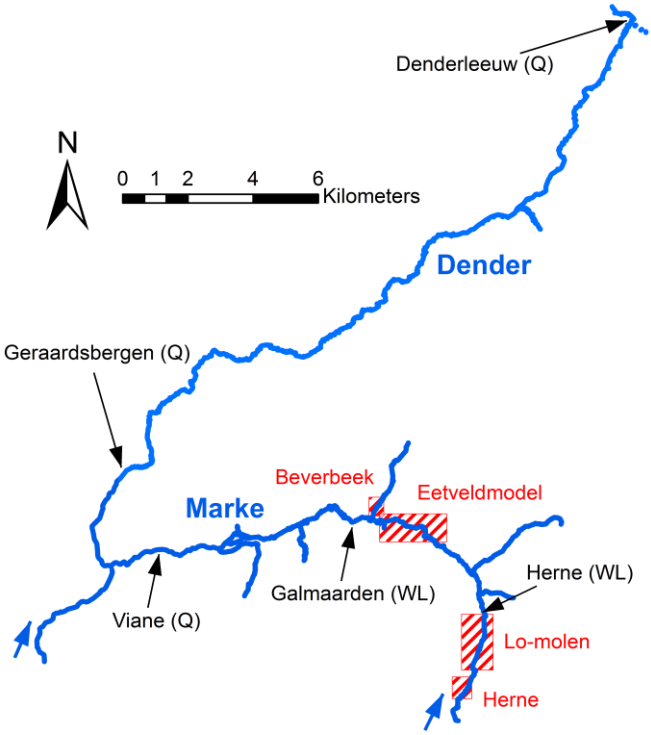


Figure 6.1: Simplified overview of part of the Marke and Dender Rivers. The selected additional retention basins are shown by the red dashed rectangles. The locations and associated states that will be analyzed are indicated as well.

The configuration of surrogate models demands for a large amount of measurements, since these models are data-driven (see §2.2.1 for a discussion on data necessity and consequences of limited data availability). Because not enough accurate measurements are available, this research employs "virtual" data sets generated by simulating different flood events in detailed full hydrodynamic models. A detailed InfoWorks RS model is available that represents the Marke River. The model was configured by the Flemish Environment Agency (VMM) and discussed extensively in Chapter 2 (§2.6.1). In addition, a MIKE11 model is

used that represents the Dender River. The model incorporates measured cross-sections every 50 to 100 meters and includes all relevant hydraulic structures. The model was calibrated carefully using gauge measurements, field surveys, maps of recently flooded areas and digital elevation models. The MIKE11 model employs a NAM model (DHI) to transform rainfall into runoff discharges. A quasi-2D approach was used to emulate the floodplains in the hydrodynamic model (see also Willems et al., 2002). This type of modelling approach is an extension of the 1D approach, in which a one dimensional river branch is linked to flood branches through link channels and lateral weirs. This method strongly reduces the computation time compared to two dimensional MIKE21 models, but delivers similar results. Both the InfoWorks RS and MIKE11 models were also used in Chapter 5 for the modelling of rating curves.

Two critical locations along the Marke River are identified that are prone to flooding and, when flooded, could result in significant economic losses. More specifically, the reduction in water depths in a floodplain in Herne and another in Galmaarden are investigated due to the selected flood countermeasures. These locations are situated just downstream of the retention basins of Herne and Eetveldmodel respectively. In addition, the effect of these measures on the receiving Dender River is quantified by analyzing the reduction of flow rates at three locations, namely (1) at the confluence of the Marke in the Dender River near the city of Viane, and along the Dender in (2) Geraardsbergen and (3) Denderleeuw. The latter two cities also suffer from floods regularly. The identified locations are indicated in Figure 6.1.

6.2.2 Investigated adaptations

To reduce the flood risk, the installation of additional retention basins together with different filling and emptying strategies are considered. Along the Marke and its tributary the Beverbeek, four locations were selected that could be used for installing retention basins (see Figure 6.1). These sites were chosen based on different criteria, including topographic elevation, land use, the potential volume of water that could be retained and ecological aspects. In total, these four locations can store nearly 1.4 million m³ water. Since some areas of the selected locations coincide with already present natural floodplains, the increase in storage capacity due to the investigated measures is somewhat lower. Any existing dikes will be altered in the simulation models to increase the current capacity, and new dikes will be incorporated to contain the flooded water.

The filling of the retention basins will be regulated by existing gated weirs on the Marke River that are installed perpendicular to the river's thalweg. By raising the gate of a more downstream structure, the water level upstream in the river will increase, thereby overtopping the retention basin's embankments and thus storing water in the floodplain. The retention basins will empty via orifices when the water level in the river decreases again. By using such filling and emptying approach, it is not necessary to build new gated weirs for three out of the four additional retention basins which reduces investment costs. The retention basin along the tributary Beverbeek (see Figure 6.1) is much smaller compared to the others, and no movable hydraulic structure was present yet to control its filling and emptying. Therefore, it was necessary to incorporate a new vertical sluice in the simulation models along this tributary. The logic controls that determine the gate levels of the weirs are configured to limit a downstream water level to a specified threshold (further referred to as "set point") to avoid floods. If the upstream retention basins reach their maximum capacity, they temporarily retain the stored water and more flow is passed downstream. To evaluate the sensitivity and effectiveness of the logic gate controls, three scenarios were implemented with high, medium and low downstream set points. The low set point is the most stringent approach, since it requires that more water must be retained upstream. This increases the risk of reaching the maximal upstream capacity before the actual peak of the storm has passed, thereby nullifying the installation of retention basins.

6.3 Conceptual model development

6.3.1 Modelling methodology and calibration current state

The conceptual modelling approach presented in Chapter 2 was used to configure the models of the Marke and Dender Rivers. The reader is referred to that chapter for a comprehensive discussion of the methodology. The development of the conceptual model of the Marke River is also treated extensively in Chapter 2 (see §2.6.2 and §2.6.3), and therefore not repeated in this section. The conceptual model of the Dender River was developed in the MSc. thesis of Vissers and Wolfs (2011) using a preliminary version of the conceptual modelling approach presented in this dissertation. The main modelling principles of this preliminary modelling approach are briefly repeated in this section. Flows in the Dender River are predominantly modelled using Transfer Functions (as discussed in

§2.3.1.4). Near the mouth of the Dender River into the Scheldt, controllable weir-type equations are used with a simplified implementation of the gate rules to ensure tidal influences and gate movements are (in an approximate manner) incorporated. The river-floodplain fluxes are calculated using fixed weir-type equations. The flow to and from floodplains is extracted from the incoming flow of the Transfer Functions to ensure the water balance is closed in the river. Most water levels in the river are determined using simple rating curves, while the stages in floodplains and in the river near the mouth are calculated using hypsometric curves (both types of calculations are discussed more extensively in §2.3.3.1). The conceptual model was calibrated and validated using the historical events of January-February 1995 and November 2010, and synthetic runoff events with 25 and 100 year return periods, derived by means of an extreme-value analysis applied to a long-term time series of rainfall-runoff discharges covering the period of 1967-2003. The composite hydrograph method (Vaes et al., 2000) has been applied to derive the synthetic events. This conceptual model was also used for flood probability mapping, which is discussed in Chapter 7.

The main findings and results of the conceptual models of the Marke and Dender River at the locations that are most relevant for this case study are treated below. The Nash-Sutcliffe efficiency values (NSE, Nash and Sutcliffe, 1970; see also Eq. 2.7) of water level and discharge predictions at the selected locations of interest (listed in §6.2.1) are summarized in Table 6.1. These NSE values were calculated based on simulation results obtained after linking all model elements in each surrogate model. Hence, deviations and inaccuracies can propagate through the system and affect predictions elsewhere. An NSE value of unity implies a perfect agreement between the predictions of the detailed full hydrodynamic model and the created surrogate model. A value of zero means that the surrogate model predictions are as accurate as using the mean of the full hydrodynamic model simulation results for all time steps. It is clear that the conceptual model manages to emulate the detailed model very accurately. The NSE values for the discharge near Galmaarden drop to approximately 0.90, which is mainly caused by small fluctuations in the gate operations of the hydraulic structure compared to the gate movement in the detailed InfoWorks RS model. Figure 6.2 compares the simulation results of the conceptual model of the Marke for the 'January 2011' event with the results of the hydrodynamic InfoWorks RS model for the selected locations.

The used conceptual model of the Marke River does not use the new discrete calculation scheme with adaptive time step (see §2.4 for a comprehensive discussion) as it was not fully developed yet when this case study was performed. Hence, the conceptual model employed in this case study differs slightly from the model presented in Chapter 2. To avoid instabilities, it was necessary to limit the time step to merely 10 seconds for all calculation nodes in this case study, resulting in a calculation time of 9.5 seconds to simulate a one-month event on an i7 processor using a single core at 3.40 GHz. As discussed in §2.6.3.3, the calculation time could have been reduced to just 1.12 seconds with the calculation scheme with adaptive time step, while retaining stable simulation solutions. The additional calculation time caused by the simpler calculation scheme with fixed time step (as used in this study) did not pose major issues in this analysis. The conceptual model of the Dender River uses a completely different calculation scheme, still written in MATLAB code. Because the system responds much smoother and slower than the Marke River, a time step of one hour could be used. The simulation time of a one month event amounts a few seconds.

Finally, both surrogate models were coupled. The conceptual model of the Marke also covers the most upstream part of the Dender up to Geraardsbergen to ensure that the complex interactions at the confluence of both rivers were captured. Since there were no significant backwater effects noticed in the flow simulation results of the detailed MIKE11 model over the weir of Geraardsbergen, which is situated a few kilometers downstream of the confluence of the Dender and Marke Rivers, it was possible to link the two models using a simple piecewise linear discharge-volume relationship. The most upstream part of the Dender River and the entire Marke River are simulated with a time step of 10 seconds, while the remaining part of the Dender is simulated using an hourly time step.

Table 6.1: Nash-Sutcliffe efficiencies (NSE) for water depth and discharge predictions at the locations of interest. Events used for calibration and validation are indicated by the letters "C" and "V" respectively.

	Water depth		**Discharge**				
	Herne	Galmaarden	Herne	Galmaarden	Viane	Geraardsbergen	Denderleeuw
Marke: Current state							
Mar 2008 (C)	n.a.	0.988	0.959	0.867	0.901	0.990	n.a.
Nov 2010 (C)	0.992	0.977	0.981	0.911	0.977	0.960	n.a.
Jan 2011 (V)	0.955	0.987	0.955	0.900	0.944	0.942	n.a.
Marke: With adaptations							
Nov 2010 high S.P. (C)	0.984	0.877	0.967	0.967	0.967	0.959	n.a.
Nov 2010 medium S.P. (C)	0.707	0.903	0.915	0.915	0.915	0.958	n.a.
Nov 2010 low S.P. (V)	0.946	0.906	0.917	0.917	0.917	0.960	n.a.
Jan 2011 high S.P. (C)	0.904	-0.417	0.956	0.956	0.956	0.936	n.a.
Jan 2011 medium S.P. (C)	0.871	0.920	0.945	0.945	0.945	0.943	n.a.
Jan 2011 low S.P. (V)	0.604	0.546	0.839	0.839	0.839	0.940	n.a.
Dender							
T25 (C)	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	0.91
Jan 1995 (V)	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	0.94
Nov 2010 (V)	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	0.96

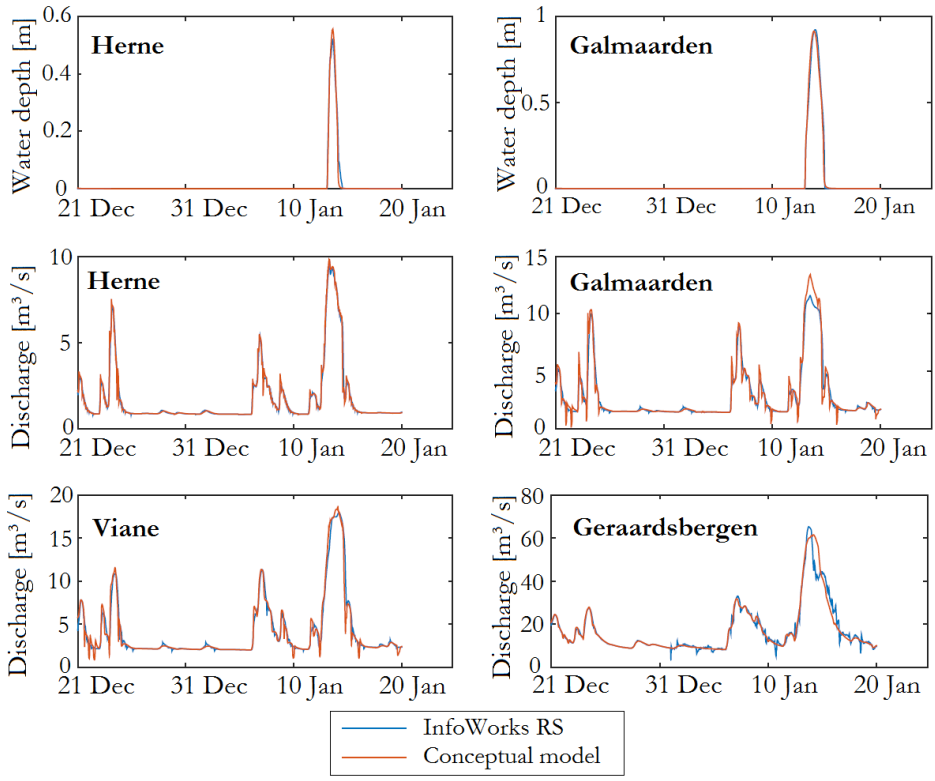


Figure 6.2: Validation results of the surrogate model of the Marke River in its current state for the January 2011 event versus the emulated simulation results of the detailed InfoWorks RS model.

6.3.2 Recalibration to include the selected adaptations

The installation of additional retention basins along the Marke discussed in §6.2.2 entails the adjustment of dike levels, thereby affecting flows, water surface profiles and the storage capacity of multiple areas. To ensure that the surrogate models can predict the flow and water levels accurately after installation of the retention basins, the Marke model is partially recalibrated. Indeed, the conceptual model is data-driven and is not designed to predict the river's states after such fundamental adjustments. Therefore, several modifications were made in the detailed InfoWorks RS model of the Marke and the storms of November 2010 and January 2011 were simulated for the three gate level configurations described in §6.2.2. These simulation results were then used to recalibrate and validate parts

of the conceptual model of the Marke River. No adaptations were made to the Dender River. Hence, it was not necessary to recalibrate this model.

The NSE-values for the flow predictions are comparable to those of the conceptual model representing the current state (see Table 6.1). The performance values for the water depth predictions in the examined floodplains of Herne and Galmaarden fluctuate however, in particular for the January 2011 storm. The spilled volumes to these floodplains are very small during this event. Predictions of water depths in floodplains are therefore very sensitive to slight deviations in the river water level forecasts. These deviations make the difference between overtopping of the dikes and hence filling the floodplains, or remaining just under the dike level. Since the volumes causing these anomalies in water depth predictions are small, they are considered insignificant.

6.4 Results

To quantify the effect of the selected additional retention basins and gate operations on floods, the storm of November 2010 and a long-term rainfall series of 36 years (1 January 1967 – 31 January 2003) was simulated with the coupled Dender-Marke surrogate model. Each period was simulated four times: once with the model reflecting the current state of both rivers, and three times with the model that includes the additional retention basins and related modifications, but each time with a different configuration of gate level controls.

6.4.1 Storm of November 2010

Table 6.2 shows the change of the peak water levels and discharges at the selected flood prone locations due to the installation of additional retention basins and the newly configured gate controls. The flood water depth near Herne is reduced by merely 15% (just over 10 cm in absolute height). The additional retention basins are much more effective in reducing the flood water levels near the city of Galmaarden, with a reduction up to 50% (i.e. almost 60 centimeters) for the storm of November 2010. The effects of the three implemented gate control configurations are comparable. The reduction of the flow of the Marke River near Viane, which is just upstream of the confluence with the Dender River, amounts to nearly 13%. Due to the large incoming flow of the Dender from more upstream situated areas, the effect of the retention basins on the flow is

reduced to 3% downstream of the confluence, and even to less than 1% further downstream. Hence, the effect of the retention basins along the Marke River on the flow of the Dender River is negligible.

Table 6.2: Change of the water level and discharge peaks due to the installation of additional retention basins and alternative gate controls for the event of November 2010.

		High S.P.	**Medium S.P.**	**Low S.P.**
Water depth changes	Herne	-13%	-16%	-14%
	Galmaarden	-47%	-50%	-46%
Discharge changes	Herne	+16%	-3%	+1%
	Galmaarden	-14%	-16%	-14%
	Viane	-11%	-15%	-12%
	Geraardsbergen	-2%	-3%	-3%
	Denderleeuw	0%	-1%	-1%

To further examine the operation of the retention basins, the relevant states around the site of Galmaarden during the peak of the storm of November 2010 are shown in Figure 6.3. The currently applied gate configuration aims to limit an upstream water level to a specified threshold value. Hence, during the peak flow, the gate is lowered to its minimal position to unburden the upstream system. The new gate configurations with different set points target the opposite by retaining water upstream in the newly created detention basins with the aim of reducing water levels more downstream. As expected, the gate level in the scenario with lowest downstream set point gate control increases the most rapidly. Eventually, the gate levels in all three scenarios reach their maximum positions. This value is imposed by the limitations of the present structure. To limit the investment cost of the adaptations, the current hydraulic structures remain unchanged as discussed in §6.2.2. In this specific case, this limitation probably does not have a significant effect on the flood water levels, since the stored volume of water in the upstream retention basin reaches almost its maximum capacity of 750 000 m³ as can be seen in Figure 6.3, although this maximum is attained later due to the gate limitations. Further research could focus on the added effect of relaxing the gate movement restrictions.

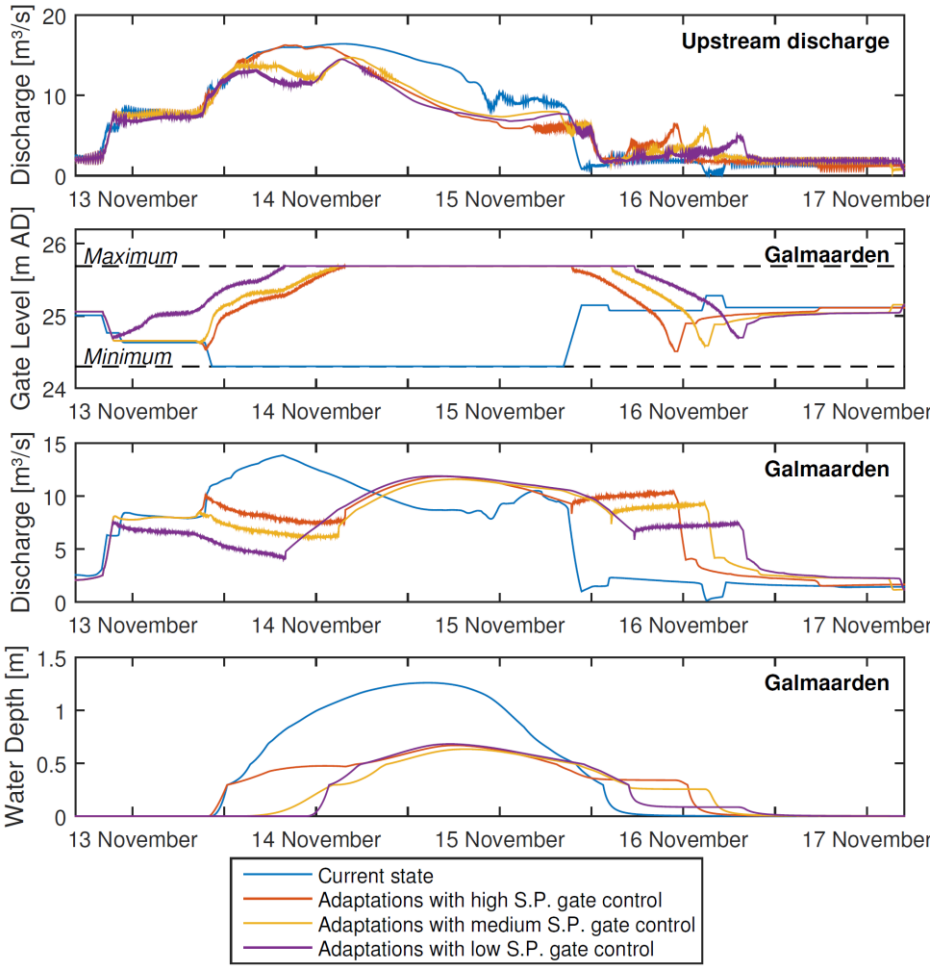


Figure 6.3: The effect of the installation of additional retention basins and gate control configurations with different set points (S.P.) on several river states near the city of Galmaarden for the peak of the November 2010 storm. The discharge and gate level states of Galmaarden were analyzed just downstream of the Eetveldmodel retention basin.

6.4.2 Long term simulation and statistical analysis

Usually, flood management plans are developed in an event-based manner by simulating peak inflows with return periods prescribed by regulations (e.g. Lammersen et al., 2002). Such approach encompasses statistical processing of the model inputs to yield the necessary design inflows. The most fundamental

problem associated with this approach is the implicit assumption that an input event with a certain probability of occurrence results in an output with the same probability. This is however not the case due to the non-linear character of the system. In addition, this technique is subjected to several large uncertainties (Parmet et al., 1999), such as (1) the use of a non-homogeneous discharge record due to climatic variability, or changes to the river system or hydrology and (2) the choice for a specific extreme-value distribution during the statistical processing. Therefore, it is preferable to use a time series-based approach, which entails the simulation of the entire available time series of rainfall observations. This approach has the advantage of avoiding a statistical preprocessing of the inputs and, perhaps even more important, it takes meteorological conditions and the catchment response explicitly into account. After such long-term simulation, a statistical processing is performed to yield information on the return periods of system states.

This study uses the surrogate models to perform such long term simulation. The necessity of using surrogate models becomes clear when the calculation times are compared to full hydrodynamic simulations: the calculation time for simulating the 36-year time series amounts a little more than one hour using the surrogate model, while the detailed models would require more than 3 weeks. Next, a peak-over-threshold analysis was performed on the simulation results using the WETSPRO tool (Willems, 2009) to select independent peak water depths and discharges. This process was performed for the current state and the three investigated adaptation scenarios. This resulted in the quantification of the water depth and discharge reduction at the selected locations versus the empirical return period (see Figure 6.4).

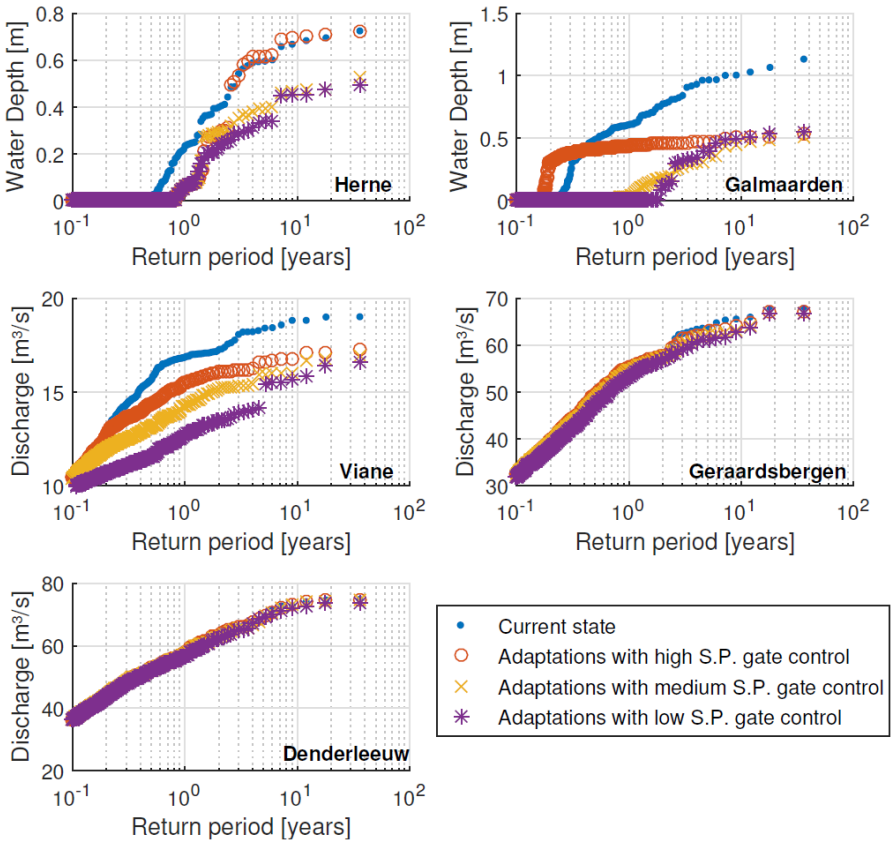


Figure 6.4: The effect of the selected measures on water depths and discharges for different return periods.

The results are in line with those of the November 2010 storm: the water depth near the city of Galmaarden is reduced effectively, the discharge at Viane is lessened by circa 10 to 20% depending on the return period, and the flow in the Dender River is hardly affected compared to the current state. In contrast to the simulations of the November 2010 storm, however, there is a clear difference between the three different gate controls. The gate controls with medium and low set points respond more rapidly to peak storms, while the high set point control resembles more the current state. The gate of the retention basin near Herne in the high set point gate control in particular moves almost synchronously with that of the current state, resulting in the same maximal flood levels. The peak flow of the storm of November 2010 surpasses those of the long term simulation. This extreme incoming flow during the storm of November 2010 forced the gate to

rise, which resulted in a comparable decrease of the maximum flood level at Herne as the two other gate control scenarios (see Table 6.2). The peak flows in the long-term simulation were not extreme enough to induce gate movement in the high set point control however. For return periods lower than one year, there are deviations noticeable between the current and adaptation scenarios for the water depth predictions near Herne. Although the gated weir moves almost identical in all simulations for these lower return periods, the water depth prediction in the investigated floodplain is divergent. This is due to the use of other model structures and parameter sets in the adaptation scenarios after recalibration (see §6.3.2). This leads to minor deviations in the water level predictions in the river. The system responds very sensitive to such deviations when this water level approximates the dike level as discussed in §6.3.2. However, the spilled volumes associated with these return periods are very low, and hence this uncertainty does not have a significant effect on the outcome of the analysis.

6.5 Discussion and conclusions

The current evolution towards integrated modelling and the broadening of research questions, such as optimization and accounting for uncertainties, necessitate the use of fast mathematical models as described in the introductory chapter of this dissertation. This case study illustrates the use of such computationally efficient models, based on the developed conceptual modelling approach described in detail in Chapter 2, for a practical application in flood management. The effect of installing additional retention basins is quantified and different gate control configurations are assessed by simulating long-term rainfall series in surrogate models that comprise the rivers Marke and Dender in Belgium. Next, a statistical analysis yields the reduction of water depths and flows at critical locations along both rivers versus empirical return periods.

The results show that the additional retention basins can lower the water depth with up to 60 centimeters near the city of Galmaarden for larger return periods, while the effect on the Dender River is insignificant. The analysis also demonstrates the importance of carefully selected logic controls for gated weirs that regulate the filling and emptying of such retention basins. Improper gate controls do not use the capacity of the system optimal.

It is clear that the surrogate models play an essential role in this analysis. They can simulate an event several hundred times faster than the conventional detailed full hydrodynamic models. This makes it possible to quickly assess different investment strategies statistically as demonstrated in this study, but even allows for optimization of management problems at catchment scale. Numerous alternative strategies and adaptations can be statistically assessed in an acceptable time frame using the proposed conceptual models. For instance, optimizing the most important hydraulic parameters of retention basins, such as the capacity and emptying or filling strategies, can take diverse criteria and factors into account, such as land use, topographic elevation, climate change, building costs and remaining potential flood damage in the entire river basin for different return periods. It is crucial to use long term rainfall series during such optimization instead of design storms, since only the former can consider antecedent conditions correctly.

7

Application 2: Flood probability mapping by means of conceptual modelling

7.1 Introduction

This chapter uses conceptual models to generate flood probability maps for the Dender River. This river is part of the operational flood forecasting system of Flanders Hydraulics Research. The forecasting system can provide confidence intervals of the forecasted water levels and exceedance probabilities of predefined alert and alarm levels at gauged locations. A conceptual model was used to transform these probability density functions of the water levels at the gauged locations to ungauged locations and finally to water levels in the floodplains. Fast simulation models are required to ensure that different quantiles can be calculated in real time, followed by the generation of the flood probability maps.

The first section gives a brief introduction of different river flood forecasting systems. Section 7.3 describes the case study area and flood forecasting system operated by Flanders Hydraulics Research. Section 7.4 discusses briefly the methodology for quantifying uncertainties, while section 7.5 provides details on the used conceptual model of the Dender. Section 7.6 describes the developed methodology for flood mapping that can be used in combination with conceptual models. Finally, the results are presented. Flood probability maps are generated

by the conceptual model of the floodplain near Overboelare along the Dender River at different times for the storm of November 2010, and compared to a map that indicates the measured flooded area of this event.

This chapter is based on the following publication:

> WOLFS, V., VAN STEENBERGEN, N., WILLEMS, P. (2012). FLOOD PROBABILITY MAPPING BY MEANS OF CONCEPTUAL MODELING. IN MUÑOZ, R. (ED.), RIVER FLOW 2012: VOL. 2. INTERNATIONAL CONFERENCE ON FLUVIAL HYDRAULICS. COSTA RICA, 5-7 SEPT. 2012 (PP. 1081-1085). LONDON: CRC PRESS, TAYLOR & FRANCIS GROUP.

7.2 River flood forecasting systems

River flood forecasting systems become widely used for real time prediction of runoff and river discharges, river water levels and inundation variables along floodplains. When these systems are linked to a GIS interface, flood maps can be produced and updated in real time. These flood maps provide a valuable source of information for water managers and emergency units during flood events. However, they do not provide information on the uncertainty of the predicted inundation levels. The uncertainty in the water levels in the floodplains typically originates from the uncertainty on the input, the model parameters and the model schematization of the flood forecasting system.

In the recent past several methods have been developed to quantify this uncertainty, like the GLUE method by Beven & Binley (1992), the Bayesian Model Averaging by Raftery (1993), the Bayesian Forecast System by Krzystofowicz (1999), the Model Conditional Processor by Todini (2008), the Quantile Regression technique by Weerts et al. (2011) and the Non-parametric Data-based Approach by Van Steenbergen et al. (2012). Each of these methods are able to quantify the uncertainty in the flood forecasts at gauged locations along modeled river stretches. The listed methods have their advantages and disadvantages, which are discussed in Coccia & Todini (2010) and Van Steenbergen et al. (2012). In this chapter, the Non-parametric Data-based Approach has been applied to quantify the uncertainty.

Next to the quantification of the uncertainty in the flood forecasts, also the careful communication and visualization of this uncertainty has to be addressed

(Kloprogge et al., 2007). One of the presentation forms are probabilistic flood maps, which show the probability of flooding for areas in the vicinity of the river. This kind of maps could help the emergency units to focus their actions on the areas with the highest flood probability. In addition different flood management strategies could be calculated in real time and the strategy which offers the lowest flood probabilities could be chosen. Given that these flood maps are produced and used during critical periods, the method and software that calculate and map the flood uncertainty have to be robust and with limited calculation time.

7.3 Case study

7.3.1 Dender River

The Dender is a strongly channelized river controlled by 8 hydraulic gates and sluices (for navigation and flood control). It has a downstream tidal influence from the river Scheldt. The Dender basin has a total area of 1384 km² of which 708 km² lies in Flanders. The river flow is strongly affected by rainfall over the upstream basin (see Figure 1.4 in the introductory chapter of the dissertation for an overview of the catchment). The Dender has an assymetric valley; the West side is relatively flat, while the East side lies in an undulating landscape with narrow, rather strongly incised river valleys. The dominant soil type is loam. The Dender basin has a high urbanization degree of 30%, concentrated in the North of the basin. The South of the basin is characterized by arable land. Larger historical floods have occurred in 1993, 1995, 1999, 2002 and 2010. Whereas the average discharge of the Dender at the upstream gauging station of Overboelare is 5 m³/s, during the recent flood period of November 2010 the discharge rose to more than 100 m³/s at this location, which resulted in severe floods.

7.3.2 Flood forecasting system

The food forecasting system used in this study is operationally run by Flanders Hydraulics Research and provides forecasts of water levels and discharges of the navigable rivers in the Flanders region of Belgium. The system contains a combination of catchment hydrological and river hydrodynamic models and a data assimilation module for real time updating. The hydrological models are lumped conceptual rainfall-runoff models implemented in the NAM module of the MIKE11 software of DHI Water & Environment (DHI, 2007a,b; Madsen,

2000). Also the hydrodynamic models are implemented in the MIKE11 river modeling system.

The hydrological models of the gauged basins are calibrated using long term (>5 year) rainfall and evapotranspiration series. The model parameters of the ungauged basins are derived from interpolation between the surrounding gauged basins.

The hydrodynamic model of the Dender River has been calibrated and validated based on recent flood events, making use of available observations of water levels and discharges along the river stretches. This hydrodynamic model was already used in the previous chapter (see §6.3.1).

The models are run several times a day. Each simulation period takes four days, consisting of a two day hindcast period and a two day forecast period. The hindcast period allows the data assimilation module to adjust the simulation results with the available observations. The data assimilation does not only create optimal initial conditions, but also allows adjusting the forecast results, based on an extrapolation of the in time correlated errors during the hindcast period.

7.4 Uncertainty analysis

Different methodologies can be followed to assess the uncertainty in flood forecasts. In this study, a technique based on statistical analysis of the model error is used. Typical for techniques that are based on the statistical analysis of model error is that they try to deduct a probability density function of model forecast error from explanatory variables, such as measured or forecasted rainfall, historical model error and current forecasts (Montanari & Grossi, 2008).

The used technique is a non-parametric data-based approach, which is described in Van Steenbergen et al. (2011, 2012). The technique starts with an uncertainty analysis of the forecast residuals (differences between the forecasts and the observations). The analysis shows that the distribution of hydrological forecast residuals is not Gaussian, skewed and heteroscedastic. Therefore a non-parametric technique was proposed, which does not require the choice of an appropriate probability distribution.

The technique first divides the historical forecast residuals in discrete classes based on the forecasted water level and the lead time. One expects that the forecast predictions are less accurate for longer lead times and extreme water levels. For each combination of forecasted water level and lead time the empirical cumulative frequency distribution is derived. From this distribution different quantiles are stored in a so called 'three dimensional error matrix'. The first dimension represents the forecasted water level class, the second dimension the lead time class and the third dimension the quantile value. Based on a 3D linear interpolation in this error matrix, the discrete probability density function of a forecasted water level, with a certain lead time, can be derived.

This methodology is used in the operational flood forecasting system of Flanders Hydraulics Research and provides confidence intervals of the forecasted water levels and exceedance probabilities of predefined alert and alarm levels at gauged locations.

The disadvantage of this method is that it cannot provide uncertainty estimates at ungauged locations or for the water levels in the floodplains. Therefore, the method has been extended in order to provide flood probability maps, which give an indication of the flood probability in a spatially variable way along the floodplains. In order to transform the probability density functions of the water levels at the gauged locations to ungauged locations and finally to water levels in the floodplains, relations derived from a conceptual river model are used. The next section describes the development of this conceptual river model.

7.5 Conceptual model of the Dender River

Mathematical river models are necessary to generate flood maps. Detailed hydrodynamic models cannot be used in this study due to their prolonged calculation times. Computationally efficient models are required for performing uncertainty analyses and the generation of multiple forecasts, especially if this should be done in real-time. Instead of using the available detailed MIKE11 model, the conceptual model configured by Vissers and Wolfs (2011) is used. This conceptual model was already described and used in Chapter 6 (see §6.3.1). Figure 7.1 shows calibration and validation results respectively for the water level simulations in the Dender River in Overboelare, which is the location of interest

in this study. Figure 7.2 shows a validation result of the conceptual model for the water level in the left floodplain of the Dender River near Overboelare.

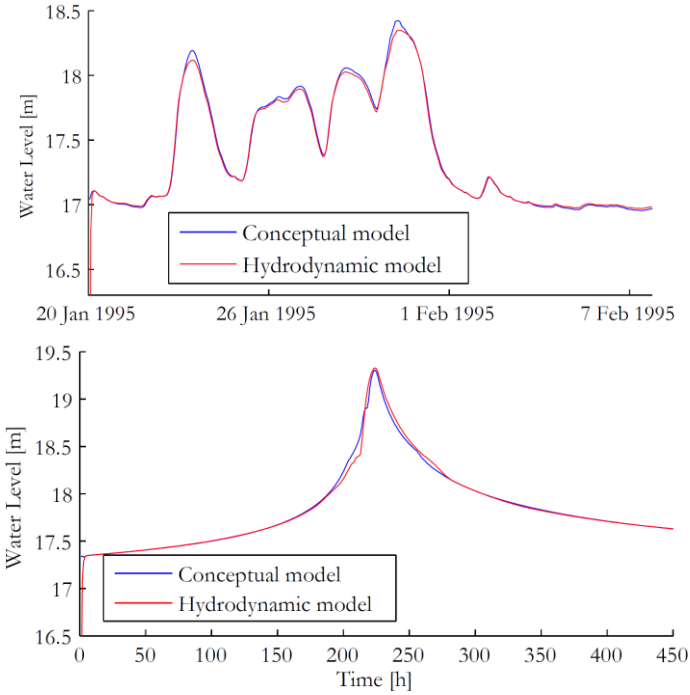


Figure 7.1: Calibration (top) and validation (bottom) result of the water level in the river Dender upstream the weir at Overboelare for the historical event of 1995 (top) and the synthetic runoff event with a 100 year return period (bottom).

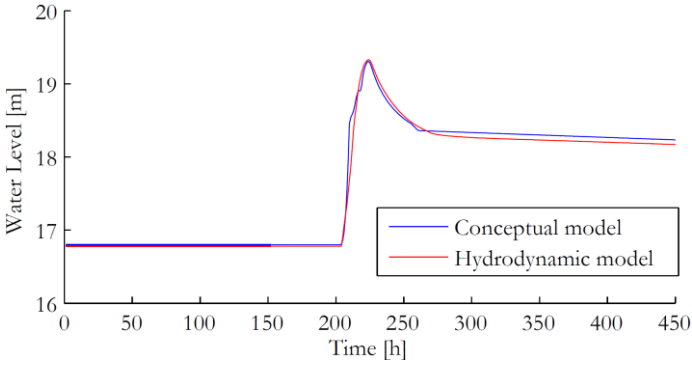


Figure 7.2: Validation results of the water level in the left floodplain of the river Dender near Overboelare for the synthetic event with a 100 year return period.

7.6 Flood mapping

In a final step the probabilistic flood maps were produced. The flood maps were generated using the technique developed in Vissers and Wolfs (2011). The technique comprises three steps. First, the digital elevation model of the investigated area is preprocessed and stored in a 2D matrix (denoted as DEM matrix) with a fixed grid size (4m x 4m in this study). Each cell of this matrix is linked to a specific floodplain in the conceptual model. Naturally, cells that do not correspond to a floodplain are not interfaced with the model and cannot inundate. Next, the conceptual model simulates an event and the obtained water levels are filled into another matrix with the same size as the preprocessed DEM matrix. The simulated water levels are compared to the level of the terrain stored in the DEM matrix: if the flood level exceeds the terrain level in a cell, that specific cell is marked as flooded. The flood depth is equal to the simulated water level minus the terrain data, and can consequently also be calculated. Finally, the generated maps can be visualized in GIS software or MATLAB. Note that the conceptual model lumps the floodplain, and only calculated one single water level for the entire flood area. The simulated water level will thus be extrapolated over the entire floodplain.

The conceptual model was used to simulate different quantiles of water levels in the floodplains, based on the quantiles of water levels at the gauged locations along the river. This was done using results of the uncertainty analysis given in §7.4. In this study the 2.5, 5, 10, 20, 50, 80, 90, 95 and 97.5% quantiles of water levels were calculated. For each of these quantiles, the associated water level in the floodplains was derived. Here the advantage of the conceptual modeling technique becomes clear, because the calculation time to derive these different water levels in the floodplains is very limited (in the range of 1-2 seconds). If the same calculations would have to be made with a hydrodynamic model, the calculation time would be significantly longer and this time is not available in flood crisis situations.

7.7 Results

The method was tested for the flood event of November 2010. The flood probability maps have been calculated for the floodplains in the region of Overboelare and are shown below for different times of forecast in Figure 7.3.

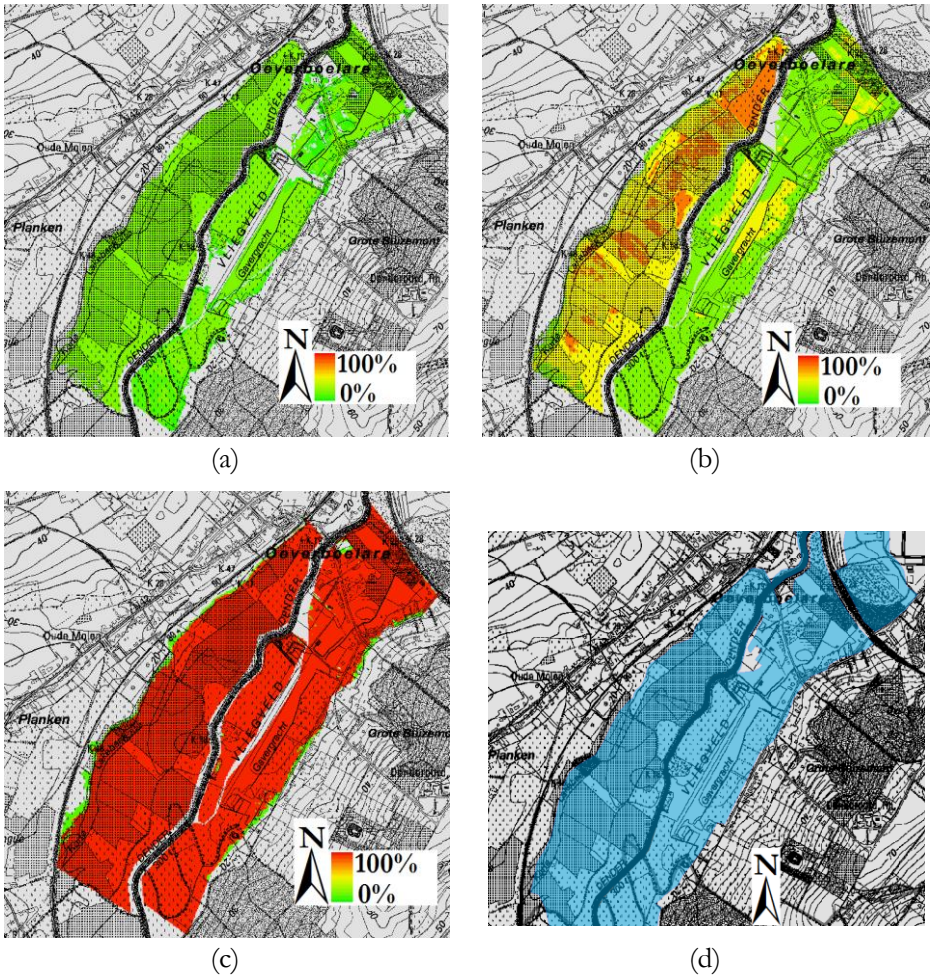


Figure 7.3: Inundation probability (%) map for the next 48h for the river Dender at Overboelare on (a) 13 November 2010 05:00; (b) 13 November 2010 10:00; (c) on 14 November 2010 01:00; and (d) Observed flood extent derived from helicopter images on the 15th November 2010.

In Figure 7.3a, a low flood probability is noticeable. Results five hours later show that the flood probability has increased significantly (Figure 7.3b). The increase in flood probability prosecuted, as can be noticed in Figure 7.3c. When the flood probability map of Figure 7.3c is compared with the observed flood extent of Figure 7.3d, a clear similarity can be observed. The locations with a high flood probability are effectively flooded, which shows the usefulness of the method.

7.8 Conclusions

This case study showed how the uncertainty in the flood forecasts can be visualized in flood probability maps. Conceptual models play a pivotal role in generating these maps. The conceptual model calculates quantiles of water levels at selected locations along the river in a very short time frame, allowing a fast production of the probabilistic flood maps. The method has shown to be very efficient for the flood event of November 2010. The observed extent of the flood matches very well with the extent shown in the flood probability map. This study demonstrates the accuracy and short calculation time of the developed conceptual modelling methodology. It also shows the applicability of the developed flood mapping module and that flood extents can be visualized precisely.

By making use of these flood probability maps in the future, water managers and flood crisis managers could focus their actions on the areas with the highest flood probability. If these maps would become available online in real time, people could check the locations of their houses and take proactive measures to reduce the personal damage.

8

Application 3: Impact analysis of CSOs on receiving river water quality

8.1 Introduction

Sewer overflows can have a devastating effect on the water quality of receiving rivers. To quantify the impact of CSOs on the river water quality, models of the river and sewer systems should be coupled. This ensures that the bilateral interactions, in which both the river and the sewer systems potentially affect each other during simulations, can be taken into account. As discussed in Section 1.2.3, the use of detailed hydrodynamic models for integrated modelling suffers from several major shortcomings and disadvantages. Therefore, the complementary use of conceptual and detailed models was advocated in Section 1.3. Such conceptual models are due to their model parsimony, computational efficiency and flexibility well suited for large scale integration. This case study illustrates the proposed generic framework for integrated modelling and water management (see Section 1.3) in practice for a specific case study in the Molse Nete catchment in Belgium. Conceptual river and sewer quantity models are calibrated based on simulation results of detailed hydrodynamic models according to the developed modelling approaches discussed in Chapters 2 and 3. The water quality of the river and sewer systems is also modelled conceptually. The conceptual water quality model was set up by PhD researcher Ingrid Keupers, according to the methodology

outlined in Keupers and Willems (2015a). Next, all conceptual models are coupled to form one simulation model. This model is then used to quantify the impact of CSOs on the river water quality via a long term simulation. Due to the computational efficiency of the conceptual models, such long term simulation can be performed in a very short time span. The impact on the river water quality is quantified by comparing the 90th and 99th percentiles with and without taking the extra discharges that enter the river through the CSOs into account.

This chapter is organized as follows. First, the study area, available data and models are discussed. Section 8.3 elaborates on the used conceptual modelling methodologies, with special attention to the integration of the two different model types. Section 8.4 summarizes the calibration and validation results of all conceptual models. Finally, Section 8.5 presents the results of the impact analysis.

This chapter is based on the following publication, worked out in collaboration with PhD researcher Ingrid Keupers who focused on the water quality aspects. The water quantity part was developed within the scope of this PhD.

> KEUPERS, I., WOLFS, V., KROLL, S., WILLEMS, P. (2015). IMPACT ANALYSIS OF SEWER OVERFLOWS ON THE RECEIVING RIVER WATER QUALITY USING AN INTEGRATED CONCEPTUAL MODEL. PROCEEDINGS OF THE 10TH CONFERENCE ON URBAN DRAINAGE MODELLING (10UDM). INTERNATIONAL CONFERENCE ON URBAN DRAINAGE MODELLING. QUÉBEC, CANADA, 20-23 SEPTEMBER 2015.

8.2 Study area, available data and detailed models

The study area is the Molse Nete catchment, located in the northeast of Belgium (Figure 8.1). The catchment has an area of 83 km^2 and is flat, with an average slope of 4‰. The Molse Nete catchment is highly urbanized with 29% of the land classified as urban and built up. Agriculture plays a main role as 48% of the land is used as grassland or cropland. The remaining 23% are natural forests, rivers and permanent wetlands. There are two waste water treatment plants (WWTP) to treat the sewage water produced in the catchment of which one discharges its effluent into the Molse Nete River and the other transports water outside of the catchment. There are four industrial sites that directly discharge their untreated water in the river, but they only emit on average 0.03 m^3/s together. Their average loading is included in the model. Besides these point sources, the rainfall runoff entering the river has a pollution load mostly originating from agricultural activities. The rainfall runoff flow is modelled using

the VHM approach developed by Willems (2014) and its pollution loads are predicted by a SENTWA model (Van Hoof, 2003; Willems et al., 2005).

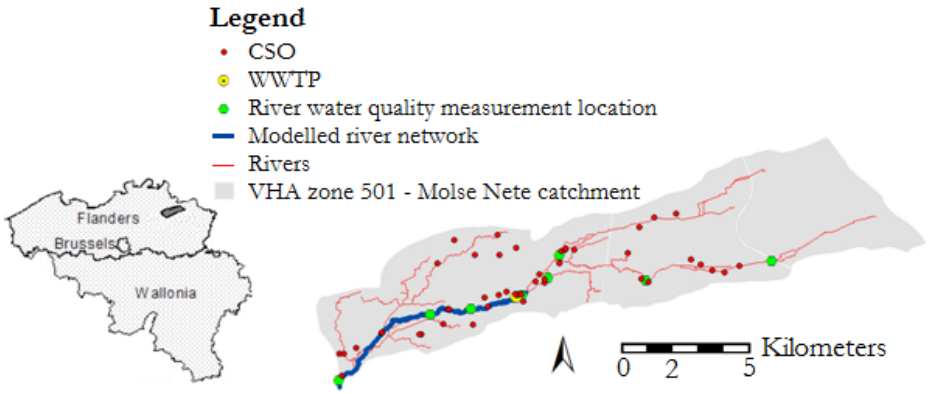


Figure 8.1: Location of the Molse Nete catchment in Belgium (left) and the modelled catchment and river network with indication of the locations of the CSOs, WWTP and water quality monitoring sites (right).

Rainfall is the driving force for both the sewer and river systems. Since the sewer system has a short response time to rainfall, it was decided to use an historical 15-minutes measured rainfall time series as input for the sewer models. It was based on the nearest rain gauge at Varendonk. Missing periods were filled based on the data from neighboring stations. For the river model, hourly rainfall records at stations of the VMM and HIC authorities were applied, corrected for the rainfall volumes at the closest, more accurate rain gauge of the Royal Meteorological Institute of Belgium (Vansteenkiste et al., 2014). There is no river flow gauging station in the catchment of the Molse Nete River, but the river model could be calibrated as part of the larger Grote Nete catchment in which four river flow gauging stations with an hourly time resolution are available. There are seven water quality monitoring points spread out over the catchment (Figure 8.1). The temporal resolution of the water quality observations are, however, limited: one sample approximately once a month, hence providing about twelve measurements per year. The water quality data can be applied for evaluating the overall water quality concentration variations in the model, but their temporal resolution is too limited to allow evaluation of the temporal dynamics at higher resolutions.

Simulation results of the detailed InfoWorks CS models of the cities of Mol and Geel were available. These models are highly detailed and include expected

changes for the coming five years. They include in total almost 8000 nodes, and have a contributing area of 5311 hectares and a population of 68749 people equivalents. The flow in the sewer systems is highly regulated: the models comprise 120 pumps and over 350 other hydraulic structures, such as weirs, orifices and flap valves.

A detailed water quality model has been set up in MIKE11 – ECO Lab which allows a continuous simulation of the hydrodynamic, advection-diffusion and water quality transformation processes. The hydrodynamic river model has been set up and validated in a previous study (Keupers & Willems, 2012). The total length of the modelled Molse Nete River is 11.5 km. Cross-section information is included every 50 meters on average. The MIKE11 river model has been extended to water quality considering oxygen, organic matter and nitrogen nutrients. The water quality state is described by seven variables, namely dissolved oxygen, BOD, ammonia, nitrate, orthophosphate (OP), particulate phosphorus (PP) and temperature using the "level-4" template of ECO Lab (DHI, 2014a). This model is considered applicable to study the effects of discharges from municipal and industrial waste and agricultural runoff (DHI, 2014b).

8.3 Conceptual water quantity and quality modelling approaches

Separate conceptual water quantity and quality models are created for the Molse Nete River and sewer systems of the cities Mol and Geel based on simulation results of detailed models. Individual calibration and validation is crucial to ensure each model has a physically correct representation of the considered subsystem (Rauch et al., 2002). The conceptual models were therefore calibrated and validated against simulation results of detailed models. No detailed model was available of the water quality of the CSOs. Due to the limited knowledge on the physical-chemical, biological and transport processes occurring in sewer systems (Bertrand-Krajewski, 2007) and limitations in sewer water quality data, such detailed models are difficult to construct, calibrate and validate. Therefore, an alternative configuration procedure was employed to set up the conceptual water quality model of the CSOs using measurements.

8.3.1 Water quantity

The newly developed conceptual water quantity modelling approach was used in this study to configure the conceptual models of the river and sewer systems. The modelling methodology for rivers (see Chapter 2) and urban drainage systems (see Chapter 3) are comprehensively described in other parts of the dissertation, and are therefore not repeated here. Note that the dynamics of the WWTP were not modelled explicitly. Instead, the influent discharges of the WWTP (which is modelled via nine pumps in the conceptual sewer model of Mol) are summed to obtain the effluent of the plant.

8.3.2 Water quality

The conceptual modeling approach for the river water quality is based on the same principles as the water quantity models. The physically based advection-diffusion and water quality transformation processes are being lumped in space and time as the river is represented by a series of interconnected reservoirs or storage cells. This lumping allows using a much larger calculation time step, which in turn leads to a significant reduction of the computation times. A detailed description of the conceptual water quality modelling approach applied for the river system can be found in Keupers and Willems (2015a). Additional information on the used conceptual water quality model is given in Keupers et al. (2015).

An ANN was configured by PhD researcher Ingrid Keupers to model the water quality of the CSOs, which was set-up according to the approach discussed in Keupers and Willems (2015b). It was not possible to apply the same conceptual modelling approach for the CSOs due to the absence of a detailed physically-based model. It was assumed that the CSO pollution load could be approximated by the measured influent concentrations at the WWTP. Only for these influent concentrations, water quality measurements are available in the sewer system, but these are limited to weekly values.

8.3.3 Integration of the conceptual models

After individual configuration of the water quantity and quality conceptual models, the different models are coupled so interactions between the river and the sewer systems can be simulated. More specifically, the simulated overflow discharges of the urban drainage systems of the cities Mol and Geel and the

estimated effluent discharges of the WWTP of Mol are considered as inputs for the water quantity and quality model of the Molse Nete River. Since many of these overflows spill into tributaries of the investigated Molse Nete River instead of directly into the modeled part of the river, an additional routing is performed on these overflow discharges. This routing is implemented by a cascade of linear reservoirs (see e.g. Viessman et al., 1989). The mean travel time is estimated based on the velocity of the water flow and the distance from the overflow to the mouth of the tributary into the Molse Nete River.

The calculation schemes of the models are written in C code (see also §2.5.1 and §3.4.5), which is not only computationally very efficient, but also ensures straightforward interfacing of the different models due to their explicit numerical calculation scheme. The interface is implemented in MATLAB to ensure user-friendliness when creating and running the models.

8.4 Calibration and validation results of the conceptual models

8.4.1 Conceptual water quantity models sewer systems Mol and Geel

Separate models were configured for the urban drainage systems of the cities Mol and Geel, because there were no interactions between both systems. The main goal of these models is to predict overflow discharges and volumes accurately.

The sewer systems of Mol and Geel are each divided into 13 storage cells. The division of a model in storage cells is case specific and depends on the intended use of the model. Throttles, pumps and other hydraulic structures usually demarcate these cells. Note that it is not necessary to partition the system at every hydraulic structure, since this would lead to extremely interlaced and complex model topologies. This typology for the conceptual sewer models was chosen after identification of the most important hydraulic structures and after assessing multiple configurations. The aim of such typology is to strike the balance between achieving model parsimony (minimize the number of calculation nodes and increase model robustness) and retaining sufficient accuracy. Figure 8.2 shows the defined conceptual model topology of the sewer system of Mol.

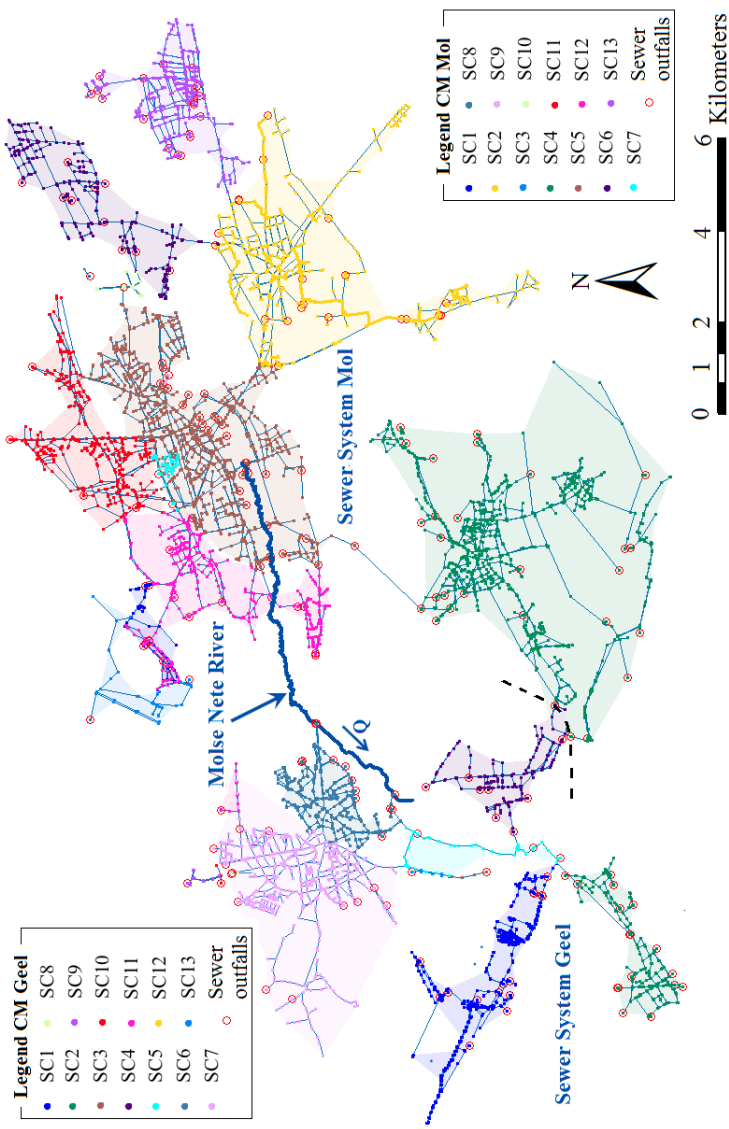


Figure 8.2: Topology of the conceptual water quantity models of the sewer systems of Mol and Geel. All outfall links of both systems are indicated.

In total, there are 80 and 31 connections where flow is estimated between the cells of Mol and Geel respectively, of which 57 and 25 lead to locations where overflow constructions are installed. Many of these connections are congregates of several neighboring individual pipes and overflow constructions, since this leads again to a more parsimonious model. These connections only aggregate flows of the same urban drainage type (i.e. combined or storm sewer). Thus, different quality parameters can still be linked to each type of overflow. Of these

overflow connections, only 24 (Mol) and 6 (Geel) lead to the investigated segment of the Molse Nete River. However, the entire system, thus including all pipes and overflow constructions, is captured in the conceptual model. This was necessary since all storage cells are (directly or indirectly) connected to each other. Because the complete sewer system is modeled, additional rainfall events can also be simulated.

Calibration was performed using simulation results of the detailed InfoWorks CS models for composite storms of 2 days with frequencies and return periods of (denoted as 'F' and 'T' in the event names) 20 (F20) and 1 (F01) per year and 5 (T05) and 20 (T20) years. Validation was performed on the events F10, F07, T02 and T10. The time step of the data is 300 s. More information on these composite storms can be found in Vaes and Berlamont (1996) and Willems (2013). Due to the large number of modeled states in the systems, it is impossible to show goodness-of-fit criteria for every simulated variable. Instead, only the three overflows of each system are presented that spill the highest volumes, since these will deteriorate the water quality most. Table 8.1 lists the relative volume error of the simulated spilled volume of the overflows in the conceptual model compared to the results of the detailed InfoWorks CS models, together with the simulated spilled volume of the InfoWorks CS results. The results for both the calibration and validation events are collected in this table. It is clear that these volumes are emulated accurately by the conceptual model, although there are some discrepancies for the events with higher frequencies. However, the magnitude of these incorrect volumes remains limited. Including additional (and preferably historical) events with lower return periods will put more emphasis on these smaller flows during model configuration and will improve model performance. The average Nash-Sutcliffe efficiency (NSE) for the volumes of the cells that are linked to investigated outfalls (and thus drive the flows to these outfalls) is 0.81 and even 0.94 for the sewer systems of Mol and Geel, respectively, for all validation events. This shows that the conceptual models have good accuracy.

Table 8.1: Overview of the calibration and validation results of the conceptual model. The relative volume errors of the spilled volumes of the 3 largest overflows in the conceptual model of Geel and Mol are compared to the results of the detailed InfoWorks CS models. The spilled volume in the InfoWorks CS simulation results is shown between brackets.

	Geel F_d		Geel F_e		Geel F_c		Mol O_d		Mol F_c		Mol O_c	
T20 (C)	1%	(30311)	7%	(26710)	-2%	(18608)	-1%	(14782)	1%	(7289)	-17%	(6689)
T10 (V)	2%	(28961)	7%	(25296)	0%	(18860)	-5%	(13326)	0%	(6185)	-5%	(5918)
T05 (C)	1%	(27841)	7%	(23625)	0%	(19094)	-9%	(11912)	0%	(5019)	3%	(5160)
T02 (V)	0%	(26600)	18%	(21421)	0%	(19365)	-10%	(9943)	1%	(3675)	7%	(4192)
F01 (C)	0 %	(26132)	10%	(16502)	0%	(19459)	-17%	(8035)	-1%	(2536)	-18%	(3379)
F07 (V)	0 %	(25574)	-20%	(5361)	0%	(19585)	-26%	(3880)	-4%	(299)	-12%	(1603)
F10 (V)	0 %	(25460)	-50%	(3131)	0%	(19610)	-31%	(3165)	-4%	(264)	-7%	(1317)
F20 (C)	1%	(25237)	-100%	(11)	0%	(19661)	-32%	(1766)	-54%	(195)	-6%	(800)

Table 8.2: Validation results of the long term simulation of the conceptual model of Mol. The deviation compares the spilled volumes (summed over the entire long term simulation) of the conceptual and detailed InfoWorks CS model results.

	MOL_ F_b	MOL_ F_d	MOL_ F_e	MOL_ I_b	MOL_ O_a	MOL_ O_c	MOL_ S_a	MOL_ S_b	SUM
Deviation	0%	12%	153%	303%	31%	-11%	13%	3%	-2%
InfoWorks CS model [m³]	193	9200	16670	158	864	396528	37331	36789	497740

Next, the long term rainfall series is simulated. Table 8.2 and Table 8.3 show again the relative volume errors and spilled volume in the InfoWorks CS model for all links for which simulation results in the detailed model were stored. Although the results of the conceptual model deviate from those of the detailed InfoWorks CS model, the summed spilled volumes in the sewer system of Mol diverge only 2% (see Table 8.2). This also shows the self-correcting nature of the conceptual model due to its mechanistic set-up: if too much flow is spilled via one link, a neighboring link connected to the same storage cell will transfer fewer mass and a new equilibrium is reached. In addition, one should note that the spilled volumes of most links in Table 8.2 are very low compared to the largest relevant outfalls listed in Table 8.1, but no simulation results of the detailed model were stored for comparison for these connections. Note that the spilled volumes were accumulated over the entire long term simulation for comparison in these tables. Naturally, it is also possible to evaluate the accuracy on separate overflow events, and to assess the simulated overflow frequency.

Table 8.3: Validation results of the long term simulation of the conceptual sewer model of Geel. The deviation compares the spilled volumes (summed over the entire long term simulation) of the conceptual and detailed InfoWorks CS model results.

	Geel_F_a	Geel_F_e
Deviation	0%	18%
InfoWorks CS model [m³]	0	431270

The calculation time of the combined conceptual sewer models amounts to 1.33 s in total to simulate a one year period using a single core of an i7 processor, including routing of the sewer overflow discharges to the investigated part of the Molse Nete River. This is much faster than the original detailed InfoWorks CS models: simulating a synthetic storm of 3 days takes approximately 24 and 5 minutes for the sewer models of Mol and Geel, respectively, using the full computational resources of an i5 processor.

8.4.2 Conceptual water quantity model Molse Nete River

The investigated segment of the Molse Nete River was divided into 9 reservoirs. Each reservoir represents a river reach of about 1 km on average. The boundaries of the reservoirs mostly coincide with bridges or culverts. The characteristics of the river stretches between these bridges do not change abruptly, making lumping

on such scales possible. Calibration of the conceptual model was based on the simulation results of the detailed MIKE11 model for the period from 8 March until 11 April 2001 with a time step of 900 seconds. Data from 12 April 2001 until 31 December 2008 were considered for validation. Since the river is not equipped with controllable hydraulic structures and the influence of variable backwater effects is very limited, transfer functions (TFs) are ideally suited to model the flow of the river. Different model structures were tested and the most appropriate TF according to the Akaike Information Criterion (AIC; Akaike, 1974) was retained at each flow calculation point. The AIC aims at striking the balance between model accuracy and model structure complexity (i.e. the number of weights in the TF). Finally, water levels are calculated in the up- and downstream parts of each reservoir using piecewise linear rating curves.

The performance of the conceptual model is exceptionally good, with NSE-values (see Eq. 2.7) for all flow and water level results exceeding 0.985 for both the calibration and validation periods. The computational time equals 0.08 s for a one year period with a time step of 300 s as opposed to 7 minutes required to simulate the same period in the MIKE11 model with a time step of 1 minute.

8.4.3 Conceptual water quality models

A detailed description of the configuration procedure of the conceptual water quality model of the Molse Nete can be found in Keupers et al. (2015) and Keupers and Willems (2015a). Calibration of the conceptual model was based on the simulation results of the detailed ECO Lab model for the period 10 January until 31 December 2001 with a time step of 1 hour. The mean error (ME), the correlation coefficient (R^2) and Nash-Sutcliffe efficiency (NSE, see Eq. 2.7) were calculated for all state variables for each reservoir. Worst simulation results were obtained for DO, which can be explained by the high temporal variability of the processes that affect the DO concentration. These simulation results show, however, close agreement between the detailed and the conceptual model results. Lowest R^2 and NSE values for all reservoirs in the validation period are 0.98 and 0.94, respectively. This is also shown in Figure 8.3 where two years of simulation results for all state variables are plotted of which the first year, 2001, was used for calibration and the second year, 2002, which was part of the validation period. The computational time is 3.6 s for a one year period when a model time step of 1 hour is considered as opposed to 10 hours required to simulate the same period in the MIKE11 model with a time step of 20 seconds. Such small time step was necessary to avoid instabilities.

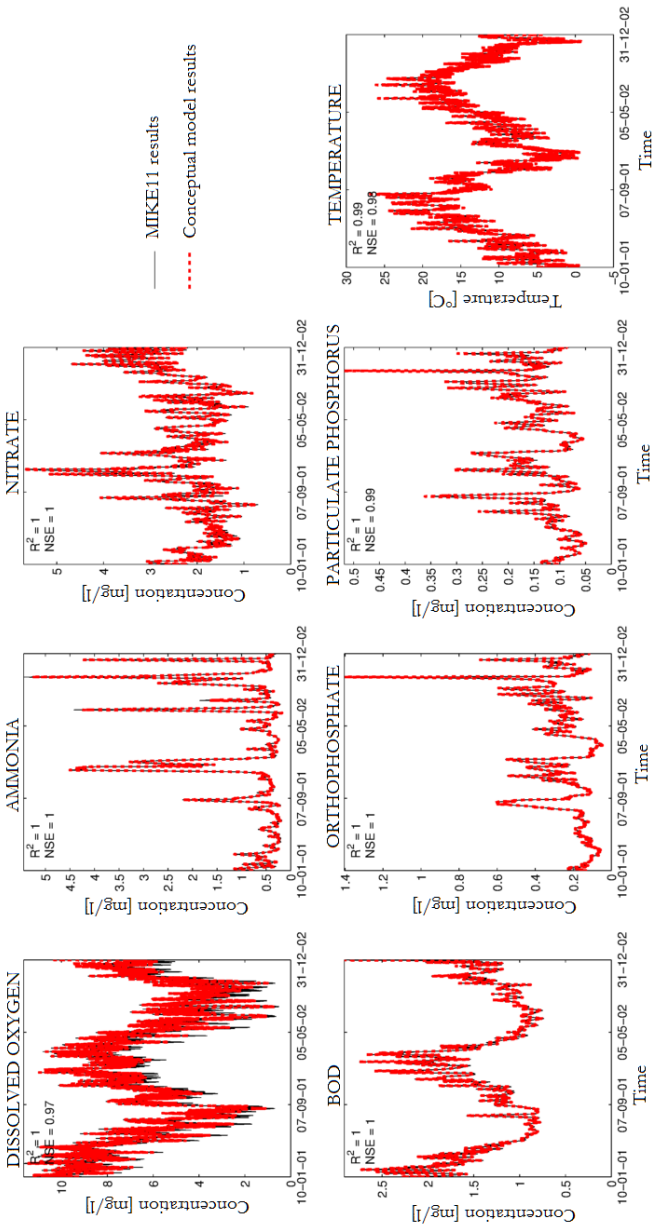


Figure 8.3: Conceptual model versus detailed model simulation results for the conceptual model output reservoir number 24 and MIKE11 results at chainage 9477, excluding the input of the CSOs

The CSO concentrations were modelled by an ANN. For each sewer system, a separate ANN was configured. The calibration and validation procedure and assumptions are described in detail in Keupers et al. (2015) and Keupers and Willems (2015b), but the main model results are briefly summarized in this

section. Table 8.4 shows that the concentrations of BOD, NH4, OP and PP can be estimated satisfactorily for the sewer system of Mol.

Table 8.4: Results water quality model sewer system Mol.

	BOD	NH4	NO3	OP	PP
# samples	103	102	39	102	102
R^2 training	0.70	0.93	-	0.94	0.91
R^2 validation	0.69	0.94	-	0.95	0.92

Table 8.5 shows that the ANN model accuracy is not very high, especially during the validation. This can probably be attributed to the large distance of the nearest available rainfall station for the WWTP of Geel, which is almost 11 km away and might not represent peak discharges over the sewer catchment accurately. This is in contrast to the WWTP of Mol where a rainfall station is situated within a radius of 7 km.

Table 8.5: Results water quality model sewer system Geel.

	BOD	NH4	NO3	OP	PP
# samples	155	133	15	132	131
R^2 training	0.66	0.93	-	0.89	0.80
R^2 validation	0.63	0.85	-	0.85	0.63

8.5 Impact analysis

The impact of the CSOs on the river quality is assessed by calculating the 90th percentile of the simulation results of the integrated model, as prescribed by the Flemish environmental regulations (VLAREM, 2010). The results show that the CSOs do not affect the conclusion on the overall compliance with the water quality standards for any of the considered water quality variables, namely DO, BOD, NH4, NO3, OP and PP. The top of Figure 8.4 illustrates this for the BOD result. This could have been expected though, since the modeled CSOs are only spilling for 8% of the time on average. However, when evaluating the impact on the extreme concentrations, i.e. 99th percentile values (Figure 8.4 bottom), an increase of almost four times the BOD concentrations can be observed at specific locations (i.e. at distance 2.3 km from upstream). The effect of these higher extreme concentrations is visible until the end of the modeled river reach.

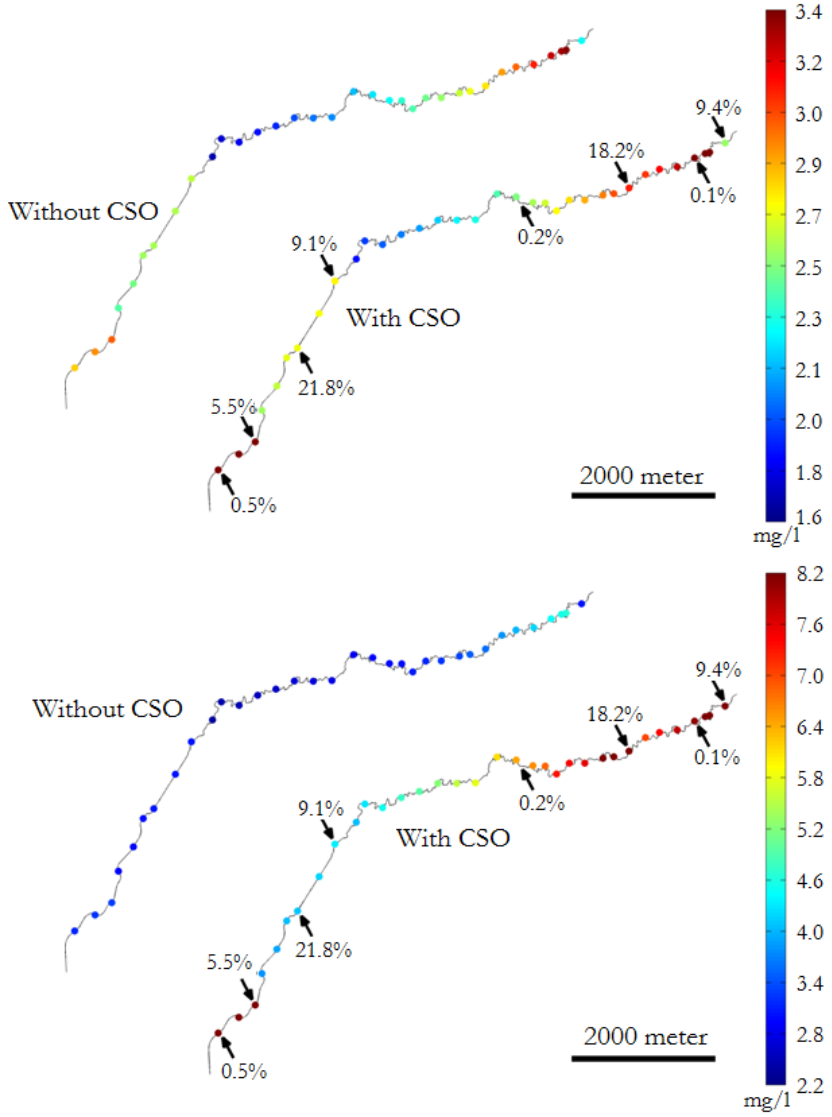


Figure 8.4: 90th percentile (top) and the 99th percentile (bottom) BOD concentration including and excluding the input of the CSOs. The arrows indicate the locations of CSOs. The percentage of the time each CSO is active is indicated (%). Note that the color scales are different of the top and bottom figures to clearly illustrate the effect of the CSOs for both percentiles.

8.6 Conclusions

In support of integrated river (catchment) management and policy making, there is a need for an integrated modeling framework. Linking the existing detailed models not only poses technical difficulties, but also does not allow for performing long term simulations due to the excessive calculation times of the complex detailed models. Such long term simulations are however required to statistically assess the impact of CSOs. Therefore, the available hydrodynamic models were converted into conceptual models according to the methodology presented in Chapter 2 (rivers) and Chapter 3 (sewer systems). Separate conceptual models were created that simulate water quantity in the river and urban drainage systems. Linking the conceptual models of the different subsystems leads to an integrated conceptual river (catchment) model that is practical because of the reduced computation times. Application to the Molse Nete catchment shows a reduction of the integrated model computation times with a factor 10 000 or higher.

As a demonstration of integrated conceptual models as decision support tools, the impact was studied of the CSOs along the Molse Nete River. It is shown that these CSOs do not have a discernible impact when evaluated for water quality status indicators based on 90^{th} percentile values. However, a strong impact on the river water quality is observed for the evaluated extreme events based on the 99^{th} percentile values. These CSO impacts may be devastating for the river ecological state, necessitating further research. It shows the importance to conduct long term model simulations and statistical analyses of the simulation results. The presented surrogate conceptual models make such simulations practically feasible.

9

General conclusions

9.1 Recapitulation

Water management is constantly evolving. Focus has long been shifting from short term objectives and local strategies to finding efficient and future-oriented solutions in a multidisciplinary context. Trends, such as the rapid urban expansion and climate change, pose new needs and challenges to water management. Risks and uncertainties have to be accounted for in order to develop sustainable strategies that remain cost-effective on the long term. Also, the need and potential benefits of a transdisciplinary and integrated approach are being acknowledged increasingly by both researchers and practitioners. Mathematical simulation models play a crucial role in support of developing water management strategies due to the intricate complexity, scope and scale of the investigated problems. These models are continuously changing to keep up with the evolutions in water management. For decades, the main objective of modellers has been to emulate reality as closely as possible, thereby increasing model detail and complexity. However, these highly detailed models do not meet the requirements for present and future water management due to several principal and insuperable shortcomings as discussed in §1.2.3. Therefore, the use of conceptual models, in addition to the detailed hydrodynamic models, is advocated increasingly in literature. Their open, lumped and simplified representation results in several essential advantages, such as a very limited calculation time, greater flexibility, enhanced integration capabilities and model parsimony, making these models ideally suited for numerous applications in water management. The literature overview revealed several fundamental deficiencies in

the gamut of conceptual modelling approaches. Existing methodologies do not meet the model requirements formulated in §1.2.1. The need for a pragmatic and effective conceptual modelling approach resulted in the definition of three main objectives.

9.1.1 Objective 1: Development of a conceptual modelling approach for river and sewer hydraulic simulations

Three subtasks were defined to achieve this objective: (i) research on different model structures, (ii) generalization, adaption and integration of the investigated model structures to form a consistent and generic conceptual modelling methodology, and (iii) investigation of the numerical model stability during simulations. The results of this research are presented in Chapters 2 to 5.

a. Research on model structures

Different model structures were inventoried, and subsequently applied on subcomponent level of the water system. Each structure was assessed for accuracy, ease and speed of the configuration, interpretability and extrapolation behaviour. First, **model structures with widespread use in hydrology** were investigated, adjusted and improved.

- **Transfer functions** (TFs), which are extensively documented and exploited in system identification theory, were linked to **linear reservoir theory**. TFs were interpreted in both the discrete and continuous domain, enabling interpretation of the structure and parameters in physically meaningful terms. Constrained real and complex pole optimization was investigated to ensure a steady state gain of unity, implying that the cumulative outflow is equal to the total incoming flow. Therefore, the water balance is closed, which is a necessary key feature for many analyses. TFs were applied for rainfall runoff modelling, flow routing through urban drainage systems and open channel flow. The TFs can be expanded with a **non-linear transformation** (NLT) to form a Hammerstein-type model as introduced by Villazon (2011) to account for floods.
- Equations that describe the flow over **(controllable) hydraulic structures** were investigated. The parameters of these higher-fidelity equations can be related directly to physical characteristics of the structure, such as a weir or sluice. Equations of movable structures also

account explicitly for gate levels, enabling easy integration in various model-based optimization applications that involve explicit consideration of hydraulic structures, such as real time flood control.

- The **static and dynamic theory**, originally developed by Vaes (1999), was investigated and generalized for sewer flow routing. This methodology was also linked to the TF theory.
- Ordinary **piecewise linear relationships** (PWL), which relate one state in the system to another, were also applied for flow and stage predictions.

In addition, a broad range of state-of-the-art **machine learning techniques and expert systems** was investigated that are not yet established in the field of hydrology, but are frequently used in other (engineering) analyses and applications.

- Extensive research was conducted on the use of **artificial neural networks** (ANNs) as an emulator of complex dynamics. Different calibration techniques were investigated to avoid overfitting, including the basic early stopping approach, Bayesian regularization and using ensemble ANNs to improve generalization. Special attention was paid to the robustness and extrapolation of the networks during simulations. ANNs were employed in studies (i) to calculate the flow from and to floodplains, but also in pipes in urban drainage systems characterized by backwater or other dynamic effects; (ii) to estimate water levels and account for rotating and quickly varying water surface profiles; and (iii) to tackle hysteretic behaviour in rating curves.
- The more complex **adaptive neuro fuzzy inference systems** (ANFIS) of the Takagi-Sugeno type were employed to calculate flows to and from floodplains. An optional algorithm to constrain target training values was investigated based on the generalized random walk theory to ensure proper ANFIS configuration. Two techniques for antecedent training were compared, namely subtractive clustering and grid partitioning.
- Different model structures were expanded with **time varying parameters** (TVP) to emulate complex phenomena, mostly using non-parametric **state dependent parameter** (SDP) techniques. Hence, parameters are made dependent of one or more system states, and subsequently parameterized using PWLs. This approach was applied to the modelling of rating curves to tackle hysteresis. The transparent

parameterization and structure allows a thorough physical interpretation of the system, as illustrated by two case studies.

- **M5' model trees** were applied successfully to model rating curves and calculate water levels. These models use classification rules in a tree-like structure with simple linear models at the leaves, making them easy to interpret.

b. Integration of model structures to develop a conceptual modelling approach

Separate methodologies were developed for rivers and urban drainage systems that share the same modelling principles, but their model topology and structures are tailored to each system. The two approaches are discussed extensively in Chapters 2 and 3. The key features are listed below.

- The two approaches have a **data-based mechanistic** nature to achieve the desired level of model accuracy, computational efficiency, interpretability in physically meaningful terms and flexibility. The methodologies are primarily aimed at nominal emulation of detailed hydrodynamic models and simplify the system on two levels, namely the model topology and the momentum equations used to estimate the system's states.
- The topology of the conceptual models was outlined and formalized based on the **storage cell concept**, meaning that the river, floodplains and urban drainage system are divided in interconnected reservoirs. The topology for sewer systems is further discretized in fluxes and subfluxes. In each reservoir, the water balance is closed explicitly during simulations. The elected structure of the conceptual model topology is highly flexible and enables the modeller to focus on the dominating processes.
- A **modular setting** was implemented so modellers can select the most appropriate model structure, depending on the behaviour of the system and the intended use. It is possible to freely connect and combine different structures in a single model, without introducing additional relationships or transformation modules. A **wide range of model structures** was incorporated based on the research from the previous subtask, ranging from more classical hydrological models such as linear reservoirs, PWLs in various settings and TFs, to machine learning

techniques and expert systems such as ANNs and M5' model trees. This wide range of structures allows capturing very complex dynamics. The different model structures were generalized and adjusted to fit in the modelling approach. The static and dynamic theory (Vaes, 1999), for instance, was adapted to comply with the modular setting and conceptual model topology. In addition, special attention is paid to the transparency, extrapolation performance and interpretability of each model structure during calibration, and the impact of limited data availability.

- The developed approach was tested on various **case studies**. The results proof that the methodology **meets the model requirements** defined in §1.2.1, and overcomes many limitations of other conceptual modelling approaches. Some interesting results are highlighted below:
    - The approach manages to model **backwater effects, and pressurized and reverse flows** in urban drainage systems accurately for the investigated case studies of Ghent, Mol and Geel. Such complex dynamics are emulated by the data-driven model structures. These structures were configured using simulation results of detailed full hydrodynamic models that included such dynamics.
    - **Flood levels, volumes and flows** of rivers and floodplains were modelled accurately for all investigated case studies, with Nash-Sutcliffe efficiencies (NSE) exceeding 90% for almost all investigated variables.
    - A **flood mapping** module was validated for the Dender River for the storm of November 2010. This study demonstrated that the conceptual models can simulate flood extents correctly.
    - **Inaccuracies induced by the conceptualization of the system are negligible** compared to the uncertainties and errors of detailed hydrodynamic models for the investigated cases.
    - The derived conceptual models are **computationally very efficient**. The calculation times were reduced with a factor from $2 \cdot 10^3$ to 10^6 compared to hydrodynamic models for the examined case studies, depending on the level of detail and chosen structures in the conceptual models.
    - The modular setting and flexible model topology allow for **lumping processes on different levels**, as demonstrated by the various case studies.

- Elements that are important for **scenario analyses**, such as dike levels and weirs, can be modelled explicitly, which is a unique feature in conceptual modelling.

c. Numerical stabilization

To prevent and resolve numerical instabilities during simulations, a **new three-phase discrete and explicit solver with self-adjusting time step is developed** that mimics the variable time step solver of continuous solutions. When activated, the solver analyzes model calculations constantly during simulations to detect erratic behaviour. The instability detection criterion is based on the size of fluctuations of the simulated states around a linear trend line. If this criterion surpasses a specified threshold, the solution of a storage cell is classified as unstable and the solution is rejected. Next, the calculation of this cell's states is reiterated with smaller time steps. The solver is particularly efficient if the model comprises one or more instability-prone cells that require smaller calculation time steps than the majority of cells, since the time step can vary both in space (i.e. different for each storage cell) and time.

9.1.2 Objective 2: Development of a semi-automatic tool for conceptual model configuration

A semi-automatic software tool, named Conceptual Model Developer (CMD), was developed that incorporates the new conceptual modelling approach. The tool guides the user in a step-wise manner through the entire configuration process. Separate tools were developed for rivers and urban drainage systems, allowing easier software maintenance, expansion and future implementation in other systems. The features and functionality of the tools are discussed in Chapters 2 and 3 and summarized below:

- The CMD tools are programmed in **MATLAB** and equipped with graphical user interfaces (**GUI**s) to improve user-friendliness.
- An **interfacing with hydrodynamic software** was foreseen to import simulation results from InfoWorks and MIKE softwares. In addition, information of the model network, parameters (such as those of hydraulic structures or catchment characteristics) and logic controls (of both hydraulic structures and abstractions) can be imported from InfoWorks models. Both tools are also compatible with spreadsheet data.

- The CMD tools are designed to **handle all data** in the background: data and already calibrated model structures are stored in data bases and invoked automatically during model set-up when needed.
- Both tools **interpret the defined conceptual model topology** and "learn" how all elements of the model are interconnected. This is a crucial feature, since it allows supporting and automating the model configuration profoundly. In addition, once the model has been set-up, the tool completely automatically assembles all configured elements into one model script that simulates selected events.
- The calculation scheme is generated in **C programming language**, which runs several orders of magnitude faster than tests with Simulink and MATLAB models. The code has been **optimized** via extensive profiling to ensure maximum calculation speed and best memory usage.
- Certain **adaptations to configured model structures** are automatically performed during model set-up. For instance, TFs are solved in the discrete domain during simulations to obviate costly solutions requiring convolution. Before implementation in the calculation scheme, TFs are discretized for the correct sampling time using a zero- or first-order hold method, since the parameters of a discrete-time TF are dependent on the sampling interval. To simplify the incorporation of ANNs, their structure is converted to, depending on their dimensionality, a vector, matrix or tensor. These structures can have a variable resolution to maximize computational efficiency: erratic spaces can be incorporated with more detail, while more smooth spaces have a larger grid size.

9.1.3 Objective 3: Application of conceptual models in various water management studies

Various applications in water management were elaborated in which conceptual models played a vital role. The applications were selected carefully to ensure that the models could be evaluated in very diverse settings and for multiple objectives. Three of these applications are discussed in detail in the dissertation.

- The **impact of flood control basins was quantified** on floods along the Dender and Marke Rivers (see Chapter 6). Supplementary basins were implemented and existing storage areas were enlarged along the Marke River in the model. In addition, alternative gate regulations were evaluated. Conceptual models with a high computational efficiency were

used to perform a long term simulation, which allows accounting for antecedent conditions correctly. Reckoning with such conditions is essential in this study, since basins might not run empty between two consecutive storms, thereby reducing the original storage capacity. Due to the mechanistic setting of the modelling methodology, the investigated measures could easily be incorporated explicitly in the conceptual model. Statistical analysis of the simulation results yielded the reduction of water levels and flows at critical locations along both rivers versus the empirical return period. Future research could focus on assessing different investment strategies statistically and optimizing, amongst others, the design of retention basins and operation rules.

- Conceptual models were also used for generating **flood probability maps** along the Dender River (see Chapter 7). The model translates prediction uncertainties calculated at gauged locations using a non-parametric data-based approach (Van Steenbergen et al., 2012) to flood probability maps. Conceptual models are required due to their fast calculation times to perform all requisite simulations in real time, so emergency units can respond rapidly to threats. Validation on the basis of the storm of November 2010 demonstrates that the conceptual model can simulate flood extents very accurately.
- Finally, an **impact analysis of combined sewer overflows (CSOs) on the river water quality** was performed (see Chapter 8). This application is a demonstration of the proposed framework towards integrated water management discussed in §1.3. Conceptual models are well suited for such analysis due to their fast computation speed, model parsimony, flexibility and integration capabilities. Conceptual models of the Molse Nete River and the urban drainage systems of the cities of Mol and Geel were configured based on simulation results of hydrodynamic models. These models were integrated and coupled to conceptual water quality models. Finally, a long term simulation and a statistical analysis of the results were performed.

In addition, the use of conceptual models was crucial in other studies that were not adopted in this dissertation. One study focuses on model-based real time control of hydraulic structures to minimize flood damage in the Dender catchment (see Van den Zegel et al., 2014 and Vermuyten et al., 2014). The conceptual models are used for model predictive control (MPC) combined with genetic algorithms to run a huge number of simulations in real time, after which

the best control strategy is retained and applied. The explicit incorporation of hydraulic structures in the conceptual models facilitates the model integration in the MPC framework. A second interesting study quantified the impact of source control and end-of-pipe solutions for river and urban floods (see De Vleeschauwer et al., 2014). An integrated river-sewer conceptual model was configured that included three different adaptation strategies. This model was used to simulate a 100-year rainfall series. The flexible and open conceptual model structure allowed to easily incorporate the different strategies and to integrate all models.

9.2 Concluding remarks

Although the developed modelling methodology meets the requirements defined in §1.2.1 and conceptual models were already successfully employed in diverse case studies, several **limitations of the conceptual modelling methodology** are summarized below.

- The **data-driven** character of the conceptual modelling methodology is both one of the greatest strengths and cause of most limitations of the approach. Due to its data-driven nature, more complex dynamics, such as variable backwater effects and rapidly varying flows in sewer pipes, can be simulated accurately in most situations. To ensure such good emulation capabilities, some of the incorporated model structures do not have a fixed prescribed form or do not depend on a predefined parameter set, but can be modified in order to capture these complex dynamics. Although the structures can be interpreted in physically meaningful terms after identification and calibration, thus being mechanistic, such purely data-driven structures themselves and their parameters can often not be linked directly to physically measurable quantities. Examples of such structures are the incorporated machine learning techniques, such as ANN, ANFIS and model trees, but also the simple PWL. Thus, the identification and calibration of the most suitable structure and parameter set is purely based on data of the calibration set. Consequently, **dynamics that are not included in the calibration data are not necessarily emulated accurately**. For instance, the optional module introduced to account for non-univocal (or rotating) water surface profiles (see §2.3.3.2) makes use of an ANN. After calibration, this module can, in most cases, precisely emulate the water surface

profile in a river stretch for the calibration data under varying conditions. However, it is possible that the conceptual model fails to emulate dynamics that are not exhibited by the calibration data. The same holds for the data-driven structures used to mimic, inter alia, pressurized flows or variable backwater effects. Consequently, a deliberate choice of calibration and validation data is very important. Naturally, the limitations on the use of data-driven model structures also allude to extrapolation beyond the calibration range. However, several measures are incorporated in the modelling methodology and configuration procedure to assess and improve the extrapolation capabilities of the derived conceptual models (see e.g. §2.6.3.2). Finally, note that not all included model structures are purely data-driven. The hydraulic structure equations that calculate flows, for instance, are directly adopted from the hydrodynamic InfoWorks models and most of their parameters can be linked to the (measurable) characteristics of the hydraulic structure itself. Thus, flows over such structures can be simulated precisely under all conditions for a given set of up- and downstream water levels.

- The data-driven nature of certain model structures also **limits the conceptual models' applicability for scenario analyses that involve fundamental changes to the water system** itself. Indeed, analyses that require alterations of the river or sewer system that cause (new) variable backwater effects or fundamentally change the system's response in any other way, necessitate model recalibration. Such model recalibration allows capturing the new dynamics of the system. The developed modelling approach allows for partial recalibration: only those segments of the model that are affected by the changes require recalibration, while others can remain unaltered. System modifications that potentially require such recalibration are, amongst others, the change of dike levels and the inclusion of new controllable hydraulic structures.
- The developed **conceptual modelling approach for sewer systems does not simulate water levels**, which limits its use for certain applications. Often, the design of (new) sewer systems involves requirements for maximal allowable water levels for certain return periods. Consequently, the modelling approach cannot be employed to check such design criteria. Also, conceptual sewer models cannot be applied for flood mapping purposes at this moment. However, extensions of the approach to meet these needs appear feasible.

- All used calibration and validation data are simulation results of detailed hydrodynamic models. In many cases, not enough accurate in-situ measurements are available to do a thorough calibration and validation of the conceptual models. The use of simulation results implies that the **conceptual models can only be as accurate as the original hydrodynamic models from which they are derived**. It is assumed that the simulation of the detailed hydrodynamic models with the de Saint-Venant equations represent reality accurately, although of course such hydrodynamic models are also a simplification of the true river or sewer system. The developed approach is not confined to a single hydrodynamic software platform or model type, but aims to be more generic. Thus, modifications to the simulation engine and more specifically the assumptions made in the used InfoWorks and MIKE hydrodynamics models will likely not affect the conceptual modelling approach fundamentally. It moreover may be useful in some cases to refine the conceptual model calibration – after the initial calibration to a detailed model – based on the available field observations.

The conceptual modelling approach was formalized and automated to a great extent by the developed software. Some elements cannot be automated fully though. In addition, the modeller must be familiar with the basics of mathematical modelling in order to use the software and to configure conceptual models properly. Therefore, **expert judgment and some manual operations remain necessary** for several aspects. The full configuration procedure for conceptual models is briefly outlined below. For each phase, the elements requiring expert knowledge or important manual operations are listed.

- The first, and arguably the most important, element involving expert judgment when setting up conceptual models is a clear **definition of the intended use of the conceptual model**. Will it be used to solely model high flows, or are emulating low flows equally important? In which parts of the model are spatially detailed simulations necessary, and which processes can be lumped in space? Defining the correct objectives and error tolerances prior to model calibration are crucial and affect the entire configuration procedure.
- Secondly, a deliberate choice of **calibration and validation data** is essential. As stressed repeatedly in this concluding chapter, part of the modelling approach is data-driven. Only dynamics exhibited by the

calibration set can be emulated accurately by purely data-driven structures. Hence, the modeller should ensure that the selected calibration set contains a sufficient amount of diverse data that represents reality closely.

- After importing the calibration and validation data, and the network characteristics of detailed models into the CMD software, **the conceptual model topology can be defined**. The conceptual model topology highly depends on the intended use, desired accuracy, available data and dynamics of the system. Therefore, it is impossible to give a conclusive list of criteria for demarcating the topology. However, based on experience, some qualitative guidelines where cells are usually demarcated for rivers and sewer systems are given in §2.2.2 and §3.3.1 respectively. The topology needs to be defined in template sheets for rivers before importing them in CMD, while the topology can be outlined directly in CMD for sewers. For the latter, a range of auxiliary tools was incorporated to assist the user. Note that all data and model structures are handled fully automatically in the background by the CMD software, and thus no manual interventions are required. Also, if part of the model topology changes during configuration, CMD automatically checks how these changes affect the model structures that are already calibrated and undertakes the most appropriate action.
- Next, the most appropriate **model structures can be identified and calibrated**. Different structures were incorporated in the modelling approach to ensure that most situations can be modelled efficiently and accurately (see Table 2.1 and §2.3 for a list of and discussion on the included model structures for rivers, and Table 3.1 and §3.3 for sewer systems). Each type of structure has its own advantages and limitations. The modeller should select the most suitable type of structure (e.g. PWL, TF, ANN, ...) to simulate a specific variable, if necessary via a trial-and-error procedure. The required algorithms to identify and calibrate the model structure of a specific elected model type are incorporated in CMD. Consequently, these actions can be performed automatically in most cases. Some model structures require additional parameters during configuration, such as the number of hidden nodes in ANNs, but all parameters have recommended default settings. Also, measures to enhance the generalization capability, such as the use of ensemble ANNs, are performed automatically. After calibration, the results of the

model structure are visualized and compared to the calibration and validation series.

- Finally, all calibrated model structures are assembled automatically in a model script by CMD to simulate events. If the **variable time step solver** (see §2.4) is used to maximize computational efficiency and obviate instabilities, it is necessary to specify parameters of this solver. Again, default settings are provided, but these can be changed on a trial-and-error basis to improve performance.

9.3 Future research

Although this research addresses the main objectives, there is always a need for improvements. The proposed additional research below focuses primarily on facilitating model identification and calibration. Other listed research topics and adaptations are recommended to gain knowledge on model performance, or ensure the applicability of the modelling approach for specific demanding applications in water management. The suggested future research can be divided into three groups.

The first group comprises additional research regarding **methodological aspects of the conceptual modelling approach**:

- Although the developed **conceptual modelling approach for urban drainage systems** is employable for the vast majority of applications in water management, two methodological extensions might increase the applicability further. Firstly, **water levels** in urban drainage systems cannot be simulated currently, which is a requisite for certain applications. It is expected that due to the various dynamic effects in pipes, relationships between stages and other states will be highly non-linear and non-univocal. It is recommended to evaluate the performance of supervised machine learning techniques, such as ANN and ANFIS which are already incorporated in the modelling approach, due to their excellent function approximation capabilities. As demonstrated in this dissertation, these structures can emulate such complex behaviour accurately. Secondly, **hydraulic structures** cannot be modelled explicitly in urban drainage systems in the current approach, mainly because they dependent on accurate water level predictions. The implementation of such equations can be a requisite for certain model-based optimization

applications of movable structures, although several alternative solutions exist in this context. Breckpot (2013), for instance, concluded that optimizing discharge variables directly instead of using hydraulic structure equations significantly simplifies the optimization problem.

- **Tides** can have a major impact on flows and stages in rivers, but also on water quality aspects (see e.g. de Brauwere et al. (2011 and 2014) for a case study on the Scheldt River and estuary). Although tides were not investigated explicitly in this doctoral research, it is assumed that tides can be represented accurately by reduced complexity models. For instance, the tidal Scheldt River acts as downstream boundary for the configured conceptual model of the Dender River (see §7.5 and Vissers and Wolfs, 2011). The tides of the Scheldt River greatly influence the dynamics of the most downstream segments of the Dender River. The gate regulations, flows and water levels of the Dender River can be emulated accurately by the conceptual model, thus accounting for these tides. In case the tides cause variable backwater effects that lead to non-univocal behaviour of the water surface profile, the optional module designed to specifically tackle such dynamics can be incorporated (see §2.3.3.2). Although this was not tested for tidal influences yet, it is expected that the model structure can easily capture such effects due to its versatility. Alternatively, one could resort to other methods, such as described in Meert et al. (2016).
- A conceptual flood mapping module was developed for river floods. A similar **system to generate maps of urban floods** is an interesting expansion of the methodology with numerous applications. The dual drainage concept (e.g. Djordjevic et al., 1999) can be exploited to capture the complex interactions between the flow in the sewer and floods at street level, and the flow between flooded zones itself.
- The use of conceptual models has great potential for real time forecasting systems. However, the forecast model must be updated on a regular basis given measurement data to ensure the conceptual model mimics reality accurately. Future research could focus on **data assimilation and model updating** techniques. Kalman filters in particular are widely used for this purpose in hydrology and other domains (e.g. Reichle et al., 2002; Clark et al., 2008, Xie and Zhang, 2010).
- This research assumes that all calibration and validation data to configure conceptual models is error-free. In addition, data availability was not an

issue due to the use of simulation results of detailed models. However, it is required to investigate the effect of **limited data availability and outliers** on the identification and calibration of the conceptual models. Furthermore, additional research should focus on the **uncertainty** of model structures and parameters, and uncertainty and error propagation during simulations.

- The developed conceptual modelling approaches are generally data-driven. Therefore, the **extrapolation behaviour, and model robustness and accuracy** for extreme conditions should be evaluated. Special attention should be paid to **effect of model adaptations on model performance** when elaborating scenario analyses, such as adjusting dike levels of the implementation of retention basins. Given the inductive character of the modelling approach, a partial recalibration can be necessary. This topic was addressed already in the dissertation (see §2.2.1, §2.6.3.2 and §5.6), but requires further research.
- Instead of implementing programmable logic controls (PLC) explicitly to determine the position of movable hydraulic structures, the use of **fuzzy controllers that emulate PLCs** can be investigated. Such approach obviates the need of predicting every system state (such as water levels or flows) invoked by PLCs, which leads eventually to enhanced model parsimony with all associated advantages. In addition, it is likely that such fuzzy controller will have a smoother response and thus results in a greater stability during simulations. Fuzzy controllers are simple to interpret since they build on a set of rules that are easy transferrable to human language, and have been used widely in engineering applications (e.g. Sarvi and Avanaki, 2015).

In addition to the methodological research above, several **software developments** for both CMD tools are expedient. None of these expansions are essential, but they improve user-friendliness, and accelerate and facilitate conceptual model configuration.

- The existing **interfacing** between the CMD tool to conceptually model rivers and the hydrodynamic **MIKE11** software can be **extended** to the same functionality as the existing interfacing with the InfoWorks software. Also interfacing with other hydrodynamic software may be useful. Furthermore, this CMD tool can be expanded with **additional GUIs** to simplify and increase the transparency of the conceptual model

configuration process, as already done for the CMD tool for urban drainage modelling.

- The **set-up procedure can be further automated**. For instance, the CMD tools could employ Input Variable Selection (IVS) methods to identify the most relevant inputs for certain model structures (see e.g. Galelli et al., 2014). Secondly, defining the model topology can be automated further. The software could, based on some essential criteria such as the designation of locations and states of interest, fully automatically propose a conceptual model topology. However, the model topology is arguably the most important element of the entire model, and it is therefore of paramount importance that any automatically created topology is critically evaluated. In addition, it will always remain necessary to, if desired, manually define or alter a model topology to incorporate expert-knowledge. It is therefore a challenging task to find the optimal balance between intensive automation and manual configuration of the topology and its structures.

In the end, this research strives towards the development of a **model-based decision support system for analyzing and optimizing integrated water management strategies**. Numerous software and methodological developments are necessary in order to establish such system and ensure a broad uptake. Three main requisite characteristics of and developments for such system are briefly discussed below.

- The system should be based on an easily adaptable **plug-and-play framework.** It must be possible to freely combine different model types in one analysis based on data and model availability, the objectives of the study and the dynamics of the system. In addition, the framework must allow for a quick implementation of model adaptations and for performing scenario investigations. Examples of possible modifications to quantity models include inserting both large and numerous small scale storage basins, changing dike levels, inserting new (movable) hydraulic structures and varying gate operation rules.
- The **interfacing** between different models is currently done manually. A transparent, efficient and expandable system is required to automate and streamline the integration of models. Developing such system necessitates various tasks. An overview of data standards used in modelling is required, and the CMD software must be made compatible

with these. Additional research is needed to determine how processes acting on different time and spatial scales can be linked efficiently. Given the requisite holistic approach in water management, the interfacing must be flexible and able to link deterministic and stochastic models, user-defined modules and third-party software packages.

- Finally, advancements are required on numerous fields to support and elaborate **transdisciplinary** analyses. An accurate and fast conceptual modelling approach is required to simulate water quality in rivers, floodplains, urban drainage systems and waste water treatment plants, that is also compatible with the in this dissertation presented modelling approaches. Given the far reaching impact of water systems, research and knowledge acquisition cannot be confined to solely water sciences alone. A multidisciplinary collaboration is needed that accounts for, amongst others, climate, society, energy, economy, ecology, agriculture and urban development.

A

Network topology spreadsheets for the CMD river software

This appendix demonstrates the use of spreadsheets to define the conceptual network topology discussed in step 2 of §2.2.2 using the developed CMD software for river systems. Note that the CMD software for sewers does not require such spreadsheets as input, but relies mainly on GUIs.

Separate spreadsheets were designed to define river reaches and floodplains, since the latter can have a very complex topology. All river reservoirs are defined in one spreadsheet, while floodplain clusters (FPC) are delineated in separate files.

Table A.1 shows part of the topology definition of river reservoir 1, corresponding to the case study discussed in Chapter 2. The important river variables that have to be emulated are enumerated (e.g. flow (Q), rainfall runoff links (RR) and water levels (WL)), together with the corresponding name under which the variable is stored in the databases. Finally, supplementary information can be given depending on the type of variable, such as the hydraulic structure type for flows. If a movable hydraulic structure with PLC is incorporated, the water levels used in such PLC do not have to be specified in spreadsheets, but are extracted automatically by the CMD software.

Table A.1: Spreadsheet of part of river reservoir 1

Reservoir #	Variable	Name	Suppl. #1	Suppl. #2
1	Q	MK47.MK46.1	Gated Weir	out
1	RR	Mopw.MDBL1.1		
1	WL	MK47	MK47_MK46	up

The definition of a FPC topology is done using four tables. The spreadsheets of FPC2 (of the case study presented in Chapter 2; see also Figure 2.9), which contains two cells and four fluxes, are discussed to illustrate the different components. First, the water levels of each cell in a FPC are defined (see Table A.2). Next, the fluxes between different cells are numbered (Table A.3). The content of each flux is defined in Table A.4. Note that multiple flows can be aggregated in one flux, and several river water levels can be defined for each flux. All these water levels will be used during model configuration. Finally, the modeller can optionally define a hydraulic structure type for a flux (Table A.5).

Table A.2: Spreadsheet of FPC2 defining the water levels in each cell.

Storage area	Name data
1	MK_RES45-46_RO
2	MK_Res_46-42_RO

Table A.3: Spreadsheet of FPC2 defining the fluxes between different cells (SA indicates storage area/cell in a FPC, while R stands for river reservoir).

	R1	R2	SA1	SA2
R1			1	
R2			2	4
SA1				
SA2			3	

Table A.4: Spreadsheet of FPC2 indicating the content of each flux and associated river water levels.

Connection #	Name Q	Name WL1	Name WL2	Name WL3
4	MK_SP.MK43.1	MK43	MK42	
-4	MK_SP.MK41.1	MK41		
2	MK_S2.MK46.1	MK46		
1	MK_S2.MK49.1	MK41	MK49	
3	MK_S2.MK_SP3			

Table A.5: Optional spreadsheet of FPC2 defining hydraulic structure types for certain fluxes.

Connection #	Type
2	Orifice

Bibliography

ABBOTT, M.B., IONESCU, F., 1976. On the numerical computation of nearly-horizontal flows. Journal of Hydraulic Research 5, pp. 97-117.

ABBOTT, M.B., REFSGAARD, J.C., 1996. Distributed Hydrological Modelling. Chapter 2: Terminology, Modelling Protocol and Classification of Hydrological Model Codes. Kluwer Academic Publishers, Springer Netherlands. 336 p.

ACHLEITNER, S., MÖDERL, M., RAUCH, W., 2007. CITY DRAIN © - An open source approach for simulation of integrated urban drainage systems. Environmental Modelling & Software, 22, pp. 1184-1195.

ACHLEITNER, S., FACH, S., EINFALT, T., RAUCH, W., 2009. Nowcasting of rainfall and of combined sewage flow in urban drainage systems. Water Science and Technology 59 (6), pp. 1145-1151.

AJMERA, T.K., GOYAL, M.K., 2012. Development of stage-discharge rating curve using model tree and neural networks: An application to Peachtree Creek in Atlanta. Expert Systems with Applications 39, pp. 5702-5710.

AKAIKE, H., 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control 19 (6), pp. 716-723.

AKINER, M.E., AKKOYUNLU, A., 2012. Modeling and forecasting river flow rate from the Melen Watershed, Turkey. Journal of Hydrology 456-457, pp. 121-129.

ALEMU, E.T., PALMER, R.N., POLEBITSKI, A., MEAKER, B., 2011. Decision support system for optimizing reservoir operations using ensemble streamflow predictions. Journal of Water Resources Planning and Management 137 (1), pp. 72-82.

AMITAI, D., AVERBUCH, A., ISRAELI, M., ITZIKOWITZ, S., 1998. Implicit-explicit parallel asynchronous solver of parabolic PDES. Journal of Scientific Computing 19 (4), pp. 1366-1404.

AQUAFIN, 2005. Hydronaut procedure, versie 6.0. Aartselaar, Belgium. [in Dutch]

ARCHER, D.R., FOWLER, H.J., 2008. Using meteorological data to forecast seasonal runoff on the River Jhelum, Pakistan. Journal of Hydrology 361 (1-2), pp. 10-23.

ARICO, C., TUCCIARELLI, T., DOTTORI, F., MARTINA, M.L.V, TODINI, E., 2008. Peak flow measurement in the Arno River by means of unsteady-state water level data analysis. Proc. International Conference on Fluvial Hydraulics. (River Flow), Cesme-Izmir, Turkey, 3-5 September 2008.

ARNBJERG-NIELSEN, K., WILLEMS, P., OLSSON, J., BEECHAM, S., PATHIRANA, A., BÜLOW GREGERSEN, I., MADSEN, H., NGUYEN, V.-T.-V., 2013. Impacts of climate change on rainfall extremes and urban drainage systems: a review. Water Science and Technology 68 (1), pp. 16-28.

ASHLEY, R.M., BALMFORTH, D.J., SAUL, A.J., BLANSKBY, J.D., 2005. Flooding in the future – predicting climate change, risks and responses in urban areas. Water Science & Technology 52 (5), pp. 265-273.

BABUŠKA, R., 1998. Fuzzy Modeling for Control. Kluwer, Boston, MS.

BACH, P.M., RAUCH, W., MIKKELSEN, P.S., MCCARTY, D.T., DELETIC, A., 2014. A critical review of integrated urban water modelling – Urban drainage and beyond. Environmental Modelling and Software 54, pp. 88-107.

BARJAS BLANCO, T., 2010. The River Demer Controlled by MPC. PhD thesis, KU Leuven, Faculty of Engineering, Belgium.

BARROS, F.J., 2007. Comparing Synchronous and Asynchronous Variable Step Size Explicit ODE Solvers: A Simulation Study. 21st International Workshop on Principles of Advanced and Distributed Simulation (PADS'07).

BATES, P.D., DE ROO, A.P.J., 2000. A simple raster-based model for flood inundation simulation. Journal of Hydrology 236, pp. 54-77.

BATES, P.D., HORRITT, M.S., FEWTRELL, T.J., 2010. A simple intertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. Journal of Hydrology 387, pp. 33-45.

BAZARTSEREN, B., HILDEBRANDT, G., HOLZ, K.-P., 2003. Short-term water level prediction using neural networks and neuro-fuzzy approach. Neurocomputing 55 (3-4), pp. 439-450.

BECHTELER, W., HARTMANN, S., OTTO, A.J., 1994. Coupling of 2D and 1D models and integration into Geographic Information Systems (GIS). In: White, W.R., Watts, J. (Eds.). Proceedings of the 2nd International Conference on River Flood Hydraulics. John Wiley and Sons Ltd., Chichester, pp. 155-166.

BEFFA, C., CONNELL, R., 2001. Two-Dimensional Flood Plain Flow. I. Model Description. Journal of Hydrologic Engineering 6, pp. 397-405.

BENNETT, N.D., CROKE, B.F.W., GUARISO, G., GUILLAUME, J.H.A., HAMILTON, S.H., JAKEMAN, A.J., MARSILI-LIBELLI, S., NEWHAM, L.T.H., NORTON, J.P., PERRIN, C., PIERCE, S.A., ROBSON, B., SEPPELT, R., VOINOV, A.A., FATH, B.D., ANDREASSIAN, V., 2013. Characterising performance of environmental models. Environmental Modelling and Software 40, pp. 1-20.

BERTRAND-KRAJEWSKI, J.-L., 2007. Stormwater pollutant loads modelling: epistemological aspects and case studies on the influence of field data sets on calibration and verification. Water Science and Technology 55(4), pp. 1–17.

BEVEN, K., 2001. Rainfall-runoff modelling: The Primer. John Wiley & Sons, LTD, Chichester. 360 p.

BEVEN, K. J., BINLEY, A. 1992. The future of distributed models : model calibration and uncertainty prediction. Hydrological Processes 6, pp. 279-298.

BEVEN, K., YOUNG, P., LEEDAL, D., 2009. Computationally efficient flood water level prediction (with uncertainty). Flood Risk Management: Research and Practice. Samuels et al. (Eds.). Taylor & Francis Group, London, ISBN 978-0-415-48507-4.

BEZDEK, J.C., 1973. Fuzzy Mathematics in Pattern Classification. PhD thesis, Applied Math. Center, Cornell University Ithaca, NY.

BHATTACHARYA, B., SOLOMATINE, D.P., 2005. Neural networks and M5 model trees in modelling water level-discharge relationship. Neurocomputing 63, pp. 381-396.

BIRKHEAD, A.L., JAMES, C.S., 1998. Synthesis of rating curves from local stage and remote discharge monitoring using nonlinear Muskingum routing. Journal of Hydrology 205, pp. 52-65.

BIRKHEAD, A.L., JAMES, C.S., 2002. Muskingum river routing with dynamic bank storage. Journal of Hydrology, 264, pp. 113-132.

BOOJI, M.J., 2005. Impact of climate change on river flooding assessed with different spatial model resolutions. Journal of Hydrology 303, pp. 176-198.

DORMAND, P., PRINCE, P.J., 1980. A family of embedded Runge-Kutta formulae. Journal of Computational and Applied Mathematics 6 (1), pp. 19-26.

BORSÁNYI, P., BENEDETTI, L., DIRCKX, G., DE KEYSER, W., MUSCHALLA, D., SOLVI, A.-M., VANDENBERGHE, V., WEYAND, M., VANROLLEGHEM, P.A., 2008. Modelling real-time control options on virtual sewer systems. Journal of Environmental Engineering and Science 7, pp. 395-410.

BOX, G.E.P., JENKINS, G.M., 1970. Time Series Analysis: Forecasting and Control. Holden-Day, San Fransisco, CA.

BOX, G.E.P., JENKINS, G.M., 1976. Time Series Analysis: Forecasting and Control, revised edn. Holden-Day, San Fransico, CA.

BOYER, M.C., 1964. Streamflow measurement. Handbook of Applied Hydrology, Chow, V.T., Ed., Chapter 15, McGraw-Hill, New York, NY.

BRADBROOK, K.F., LANE, S.N., WALLER, S.G., BATED, P.D., 2004. Two dimensional diffusion wave modelling of flood inundation using a simplified channel representation. International Journal of River Basin Management 2-3, pp. 211-223.

BRATER, E.F., KING, H.W., 1976. Handbook of Hydraulics, McGraw-Hill Book Company, New York, NY.

BRECKPOT, M., 2013. Flood Control of River Systems with Model Predictive Control. PhD thesis, KU Leuven, Faculty of Engineering, Belgium.

BRECKPOT, M., AGUDELO MANOZCA, O., MEERT, P., WILLEMS, P., DE MOOR, B., 2013. Flood control of the Demer by using Model Predictive Control. Control Engineering Practice 21 (12), pp. 1776-1787.

BREIMAN, L., FRIEDMAN, J.H., OLSHEN, R.A., STONE, C.J., 1984. Classification and Regression Trees, Wadsworth, Belmont CA., 368 p.

BRYSON, A.E., HO, Y.C., 1969. Applied Optimal Control: optimization, estimation and control. Blaisdell Publishing Company, Waltham, MAa,. 481 p.

BUJON, G., HERREMANS, L., PHAN, L., 1992. FLUPOL: A forecasting model for flow and pollutant discharge from sewerage systems during rainfall events. Water Science and Technology 25 (8), pp. 207-215.

BURGER, G., FACH, S., KINZEL, H. RAUCH, W., 2010. Parallel computing in conceptual sewer simulations. Water Science and Technology 61 (2), pp. 283-291.

BYRD, R.H., GILBERT, J.C., NOCEDAL, J., 2000. A trust region method based on interior point techniques for nonlinear programming. Mathematical Programming 89 (1), pp. 149-185.

CASTELLETTI, A., GALELLI, S., RESTELLI, M., SONCINI-SESSA, R., 2012a. Data-driven dynamic emulation modelling for the optimal management of environmental systems. Environmental Modelling & Software 34, pp. 30-43.

CASTELLETTI, A., GALELLI, S., RATTO, M., SONCINI-SESSE, R., YOUNG, P.C., 2012b. A general framework for Dynamic Emulation Modelling in environmental problems. Environmental Modelling & Software 34, pp. 5-18.

CHANG, F.J., CHANG, Y.T., 2006. Adaptive neuro-fuzzy inference system for prediction of water level in reservoir. Advances in Water Resources 29, pp. 1-10.

CHANSON, H., 2004. The Hydraulics of Open Channel Flow, 2nd edition. Elsevier Butterworth-Heinemann, Oxford, UK.

CHAUDRY, M.H., 2008. Open-Channel Flow, 2nd edition. Springer, New York, NY.

CHIANG, P.-K., 2015. Flood Control Combining Optimization Techniques with Hydrologic-Hydraulic Modelling. PhD thesis, KU Leuven, Faculty of Engineering, Belgium.

CHIANG, P.-K., WILLEMS, P., 2013. Model conceptualization procedure for river (flood) hydraulic computations: Case study of the Demer River, Belgium. Water Resources Management, 27 (12), pp. 277-289.

CHIU, S.L., 1994. Fuzzy model identification based on clustering estimation. Journal of Intelligent and Fuzzy Systems 2, pp. 267-278.

CHOCAT, B., BARRAUD, S. , THIBAULT, S. (1983). Présentation du modèle de propagation en conduit du système S.E.R.A.I.L. Colloque de modélisation des eaux pluviales,. 47-63. Montréal, Canada: Ecole Polytechnique de Montréal and US EPA.

CHOW, V.T., 1959. Open-Channel Hydraulics. McGraw-Hill Book Co., Inc., New York.

CHOW, V.T., MAIDMENT, D.R., MAYS, L.W., 1988. Applied Hydrology. McGraw-Hill, USA.

CLARK, M.P., RUPP, D.E., WOODS, R.A., ZHENG,X., IBBITT, R.P., SLATER, A.G., SCHMIDT, J., UDDSTROM, J., 2008. Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model. Advances in Water Resources 31 (10), pp. 1309-1324.

CLOKE, H.L., PAPPENBERGER, F., 2009. Ensemble flood forecasting: A review. Journal of Hydrology 375, pp. 613-626.

COCCIA, G., TODINI, E., 2010. Recent developments in predictive uncertainty assessment based on the model conditional processor approach. Hydrology and Earth System Sciences Discussions 7, pp. 9219-9270.

COLEMAN, T.F., LI, Y., 1994. On the Convergence of Reflective Newton Methods for Large-Scale Nonlinear Minimization Subject to Bounds. Mathematical Programming 67 (2), pp. 189-224.

COLEMAN, T.F., LI, Y., 1996. An Interior, Trust Region Approach for Nonlinear Minimization Subject to Bounds. SIAM Journal on Optimization 6, pp. 418-445.

CONSTANTINESCU, E.M., SANDU, A., 2007. Multirate timestepping methods for hyperbolic conservation laws. Journal of Scientific Computing 33 (3), pp. 239-278.

CUNDERLIK, M.J., 2003. Hydrologic model selection for the CFCAS project: Assessment of water resources risk and vulnerability to changing climatic conditions. Project Report I. University of Western Ontario, Canada.

CUNGE, J.A., 1969. On the subject of a flood propagation computation method (Muskingum method). Journal of Hydraulic Research 7 (2), pp. 205-230.

CUNGE, J.A., 1975. Two dimensional modeling of floodplains. In: Unsteady Flow in Open Channels. (Mahmood, K., Yevjevich, V., Ed.), Water Resources Publications, P.O. Box 303, Fort Collins, Co. USA, Vol. II, Chapter 17, pp. 705-762.

CUNGE, J.A., HOLLY, F.M., VERWEY, A., 1980. Practical Aspects of Computational River Hydraulics. Pitman, London.

DAVID, P., 2002. Engineering the State: The Huai River and Reconstruction in Nationalist China 1927-1937. Routledge. ISBN 0-415-93388-9. pp. xvii, pp. 61-70.

DAWSON, C.W., WILBY, R.L., 2001. Hydrological modelling using artificial neural networks. Progress in Physical Geography 25 (1), pp. 80-108.

DE BRAUWERE, A., DE BRYE, B., SERVAIS, P., PASSERAT, J., DELEERSNIJDER, E., 2011. Modelling Escherichia coli concentrations in the tidal Scheldt river and estuary. Water Research 45, pp. 2724-2738.

DE BRAUWERE, GOURGUE, O., DE BRYE, B., SERVAIS, P., OUATTARA, N.K., DELEERSNIJDER, E., 2014. Integrated modelling of faecal contamination in a densely populated river-sea continuum (Scheldt River and Estuary). Science of Total Environment 468-469, pp. 31-45.

DE VLEESCHAUWER, K., WEUSTENRAAD, J., WOLFS, V., WILLEMS, P., 2013. Comparison of the up- versus downstream storage on a sewer-river system: a quantitative case study. Proceedings of the 35th IAHR World Congress, Chengdu, China. 8-13 September, Art. No. A10783. Tsinghua University Press.

DE VLEESCHAUWER, K., WEUSTENRAAD, J., NOLF, C., WOLFS, V., DE MEULDER, B., SHANNON, K., WILLEMS, P., 2014. Green-blue water in the city: quantification of the impact of source control versus end-of-pipe solutions on sewer and river floods. Water Science and Technology 70 (11), pp. 1825-1837.

DEKA, P., CHANDRAMOULI, V., 2003. Fuzzy-neural network model for deriving river stage-discharge relationship. Hydrological Sciences Journal 48(2), pp. 197-209.

DEMERITT, D., CLOKE, H., PAPPENBERGER, F., THIELEN, J., BARTHOLMES, J., RAMOS, M.-H., 2007. Ensemble predictions and perceptions of risk, uncertainty, and error in flood forecasting. Environmental Hazards 7, pp. 115-127.

DHI, 2007a. Mike 11: A modelling system for rivers and channels. Reference Manual. DHI Software 2007, 516 p.

DHI, 2007b. Mike 11: A modelling system for rivers and channels. User Guide. DHI Software 2007, 460 p.

DHI, 2014a. WQ templates. Denmark: DHI Water & Environment, 54 p.

DHI, 2014b. ECOLab User guide. DHI Water & Environment, Denmark, 150 p.

DI BALDASSARRE, G., MONTANARI, A., 2009. Uncertainty in river discharge observations: a quantitative analysis, Hydrology and Earth System Sciences 13, pp. 913-921.

DI SILVIO, G., 1969. Flood wave modification along prismatic channels. Journal of Hydraulic Division, ASCE, 95 (5), pp. 1589-1614.

DIRCKX, G., SCHÜTZE, M., KROLL, S., THOEYE, C., DE GUELDRE, G., VAN DE STEENE, B., 2011. Cost-efficiency of RTC for CSO impact mitigation. Urban Water Journal 8 (6), pp. 367-377.

DISKIN, M.H., 1964. A basic study of the linearity of the rainfall-runoff process in watersheds. Ph.D. Thesis, University of Illinois, Urbana.

DISKIN, M.H., 1967. On the solution of the Musking method of flood routing equation. Journal of Hydrology 5, pp. 286-289.

DJORDJEVIC, S., PRODANOVIC, D., MAKSIMOVÍC, C., 1999. An approach to simulation of dual drainage. Water Science and Technology 39 (9), pp. 95-103.

DOBLER, C., BÜRGER, G., STÖTTER, J., 2012. Assessment of climate change impacts on flood hazard potential in the Alpine Lech watershed. Journal of Hydrology 460-461, pp. 29-39.

DOOGE, J.C.I., 1959. A general theory of the unit hydrograph. Journal of Geophysical Research 64 (1), pp. 241-256.

DOOGE, J., NAPIÓRKOWSKI, J., STRUPCZEWSKI, W., 1987. Properties of the generalized downstream channel response. Acta Geophysica Polonica, XXXV (6), pp. 967-972.

DOOGE, J., STRUPCZEWSKI, W.G., NAPIORKOWSKI, J.J., 1982. Hydrodynamic derivation of storage parameters of the Muskingum model. Journal of Hydrology 54, pp. 371-387.

DOTTO, C.B.S., BACH, P.M., ALLEN, R., WONG, T., DELETIC, A., 2014. Towards Water Sensitive Urban Precincts: Modelling Stormwater Management Opportunities. In: Proceedings of the 13th Conference on Urban Drainage, Sarawak, Malaysia, 7-12 September 2014.

DOTTORI, F., MARTINA, M.L.V., TODINI, E., 2009. A dynamic rating curve approach to indirect discharge measurement. Hydrology and Earth System Sciences 13, pp. 847-863.

DOTTORI, F., TODINI, E., 2010. Reply to Comment on "A dynamic rating curve approach to indirect discharge measurement by Dottori et al. (2009)" by Koussis (2009). Hydrology and Earth System Sciences 14, pp. 1099-1107.

DUAN, Q., SOROOSHIAN, S., GUPTA, V., 1992. Effective and Efficient Global Optimization for Conceptual Rainfall-Runoff Models. Water Resources Research 28 (4), pp. 1015-1031.

DUCHESNE, S., MAILHOT, A., DEQUIDT, E., VILLENEUVE, J.-P., 2001. Mathematical modeling of sewers under surcharge for real time control of combined sewer overflows. Urban Water 3, pp. 241-252.

DYMOND, J.R., CHRISTIAN, R., 1982. Accuracy of discharge determined from a rating curve. Hydrological Sciences Journal – Journal des Sciences Hydrologiques 27 (4), pp. 493-504.

ELMQVIST, H., MATTSON, S.E., OLSSON, H., 2002. New Methods for Hardware-in-the-Loop-Simulation of Stiff Models. In: Martin otter, hilding Elmqvist, and Peter Fritzson, eds. 2nd International Modelica Conference, Oberfpaffenhofen, Germany. March 18-19, 2002.

ELSHORBAGY, A., CORZO, G., SRINIVASULU, S., SOLOMATINE, D.P., 2010. Experimental investigation of the predictive capabilities of data driven modeling techniques in hydrology - Part 1: Concepts and methodology. Hydrology and Earth System Sciences 14, pp. 1931-1941.

EM-DAT: The OFDA/CRED International Disaster Database – www.emdat.be, Université Catholique de Louvain, Brussels (Belgium).

ESTRELA, T., QUINTAS, L., 1994. Use of a GIS in the modelling of flows on floodplains. In: White, W.R., Watts, J. (Eds.). 2nd International Conference on River Flood Hydraulics. John Wiley and Sons Ltd., Chichester, pp. 177-190.

EUROPEAN ENVIRONMENT AGENCY (EEA), 2006. Urban sprawl in Europe: The ignored challenge. Luxembourg: Office for Official Publications of the European Communities.

FABER, B.A., STEDINGER, J.R., 2001. Reservoir optimization using sampling SDP with ensemble streamflow prediction (ESP) forecasts. Journal of Hydrology 249, pp. 113-133.

FAUSETT, L., 1994. Fundamentals of Neural Networks: Architectures, Algorithms and Applications. Prentice-Hall International., Inc., USA., 461 p.

FEFFERMAN, C., 2000. Existence and smoothness of the Navier-Stokes equation. Clay Millennium Problems (http://www.claymath.org/sites/default/files/navierstokes.pdf, last accessed 01.10.2015)

FENTON, J.D., 1992. Reservoir routing. Hydrological Sciences – Journal des Sciences Hydrologiques 37 (3), pp. 233-246.

FENTON, J.D., KELLER, R.J., 2001. The calculation of streamflow from measurements of stage. Technical Report 01/6. September 2001. Cooperative Research Centre for Catchment Hydrology, Melbourne, Australia, 84 p.

FERGUSON, B., FRANTZESKAKI, N., BROWN, R., 2013. A strategic program for transitioning to a water sensitive city. Landscape and Urban Planning 117, pp. 32-45.

FLEMING, G., 1975. Computer Simulation Techniques in Hydrology. New York: Elsevier.

FORSTER, S., CHATTERJEE, C., BRONSTERT, A., 2008. Hydrodynamic simulation of the operational management of a proposed flood emergency storage area at the Middle Elbe River. River Research and Applications 24 (7), pp. 900-913.

FRANCHINI, M., BERNINI, A., BARBETTA, S., MORAMARCO, T., 2011. Forecasting discharges at the downstream end of a river reach through two simple Muskingum based procedures. Journal of Hydrology 399, pp. 335-352.

FRATINI, C.F., GELDOF, G.D., KLUCK, J., MIKKELSEN, P.S., 2012. Three points approach (3PA) for urban flood risk management: a tool to support climate change adaptation through transdisciplinary and multifunctionality. Urban Water Journal 9 (5), pp. 317-331.

FREAD, D.L., 1975. Computation of stage-discharge relationship affected by unsteady flow, Water Resources Bulletin, 11 (2), pp. 429-442.

FREAD, D.L., 1985. Channel Routing. In: M.G. Anderson and T.P. Burt (Editors), Hydrological Forecasting. Wiley, New York, NY.

FREDERICKS, J.W., LABADIE, J.W., ALTENHOFEN, J.M., 1998. Decision Support System for Conjunctive Stream-Aquifer Management. Journal of Water Resources Planning and Management 124 (2), pp. 69-78.

FRENI, G., MANNINA, GI, VIVIANI, G., 2009. Assessment of data availability influence on integrated urban drainage modelling uncertainty. Environmental Modelling & Software 24, pp. 1171-1181.

FRENI, G., MANNINA, G., 2010. Uncertainty in water quality modelling: the applicability of variance decomposition approach. Journal of Hydrology 394 (3-4), pp. 324-333.

FRENI, G., MANNINA, G., VIVIANI, G., 2010. Urban water quality modelling: a parsimonious holistic approach for a complex real case study. Water Science and Technology 61 (2), pp. 521-536.

FU, G., KHU, S.-T., BUTLER, D., 2009. Use of surrogate modelling for multiobjective optimisation of urban wastewater systems. Water Science and Technology 60 (6), pp. 1641-1647.

GALELLI, S., CASTELLETTI, A., 2013. Assessing the predictive capability of randomized tree-based ensembles in streamflow modelling. Hydrology and Earth System Sciences 17, pp. 2669-2684.

GALELLI, S., HUMPHREY, G.B., MAIER, H.R., CASTELLETTI, A., DANDY, G.C., GIBBS, M.S., 2014. An evaluation framework for input variable selection algorithms for environmental data-driven models. Environmental Modelling and Software 62, pp. 33-51.

GERGOV, G., 1971. Determination of the loop discharge rating curve for flood wave propagation. Journal of Hydraulic Research 9, pp. 309-319.

GLANTZ, M.H., 2003. Climate Affairs: A Primer. Island Press. ISBN 1-55963-919-9. 252 p.

GOODRICH, R.D., 1931. Rapid calculation of reservoir discharge. Civil Engineering 1, pp. 417-418.

GOSWAMI, M., O'CONNOR, K.M., 2007. Real-time flow forecasting in the absence of quantitative precipitation forecasts: a multi-model approach. Journal of Hydrology 334, pp. 125-140.

GOYAL, M.K., BURN, D.H., OJHA, C.S.P., 2012. Evaluation of Machine Learning Tools as a Statistical Downscaling Tool: Temperatures Projections for Multi-Stations for Thames River Basin, Canada. Theoretical and Applied Climatology Springer Netherlands, Volume 108 (3-4), pp. 519-534.

GOYAL, M.K., OJHA, C.S.P., 2012. Downscaling of Precipitation on a Lake Basin: Evaluation of Rule and Decision Tree Induction Algorithms. Hydrology Research 43 (3), pp. 215-230.

GRUM, M., LONGIN, E., LINDE, J.J., 2004. A Flexible and Extensible Open Source Tool for Urban Drainage Modelling: www.WaterAspects.org. In: Proceedings 6th International Conference on Urban Drainage Modelling (6UDM). Institute for Urban Water Management, Technische Universität Dresden, Dresden, Germany, September 15-17 2004.

GÜNTHER, M., KVAERNØ, A., RENTROP, P., 2001. Multirate partitioned Runge-Kutta methods. BIT Numerical Mathematics 41 (3), pp. 504-514.

HÄGGLUND, T., 1995. A control-loop performance monitor. Control Engineering Practice 3 (11), pp. 1543-1551.

HÄGGLUND, T., 2005. Industrial implementation of on-line performance monitoring tools. Control Engineering Practice 13 (11), pp. 1383-1390.

HAND, D.J., HEIKKI, M., PADHRAIC, S., 2001. Principles of Data Mining. The MIT press, Cambridge, MA, 546 p.

HARLEY, B.M., 1967. Linear Routing in Uniform Channels. Thesis, University College, Cork, Ireland.

HARREMOËS, P., RAUCH, W., 1996. Integrated design and analysis of drainage systems, including sewers, treatment plant and receiving waters. Journal of Hydraulic Research 34 (6), pp. 815-826.

HARRIS, G.S., 1970. Real time routing of flood hydrographs in storm sewers. Journal of the Hydraulic Division 96 (HY6), pp. 1247-1260.

HAUGER, M.B., RAUCH, W., LINDE, J.J., MIKKELSEN, P.S., 2002. Cost benefit risk – a concept for management of integrated urban wastewater systems? Water Science and Technology 45 (3), pp. 185-193.

HENDERSON, F.M., 1966. Open channel flow, Macmilliam Series in Civil Engineering, Macmilliam eds., New York, USE, 522 p.

HERSCHY, R.W., 1978. Accuracy in Hydrometry. Wiley, New York, NY, pp. 353-397.

HERSCHY, R.W., 1995. Streamflow Measurement, E & FN Spon, London, second edition, 524 p.

HIDAYAT, H., VERMEULEN, B., SASSI, M.G., TORFS, P.J.J.F., HOITINK, A.J.F., 2011. Discharge estimation in a backwater affected meandering river. Hydrology and Earth System Sciences 15, pp. 2717-2728.

HORRITT, M.S., BATES, P.D., 2001. Predicting floodplain inundation: raster-based modelling versus the finite-element approach. Hydrological Processes 15, pp. 825-842.

HUNTER, N.M., HORRITT, M.S., BATES, P.D., WILSON, M.D., WERNER, M.G.F., 2005. An adaptive time step solution for raster-based storage cell modelling of floodplain inundation. Advances in Water Resources 28, pp. 975-991.

HUNTER, N.M., BATES, P.D., HORRITT, M.S., WILSON, M.D., 2007. Simple spatially-distributed models for predicting flood inundation: A review. Geomorphology 90, pp. 208-225.

HUNTER, N.M., BATES, P.D., NEELZ, S., PENDER, G., VILLANUEVA, I., WRIGHT, N.G., LIANG, D., FALCONER, R.A., LIN, B., WALLER, S., CROSSLEY, A.J., MASON, D., 2008. Benchmarking 2D hydraulic models for urban flood simulations. Proceedings of the Institution of Civil Engineers: Water Management 161 (1), pp.13-30.

INNOVYZE, 2013. InfoWorks RS Documentation, Hydraulic Studies Report EX1296. Cardiff Perpheral Distribution Road – Butetown Link. Ref. Type: Technical Document.

INNOVYZE, 2014a. InfoWorks RS Help Documentation [Version 14.0]. Innovyze, Oxfordshire, UK.

INNOVYZE, 2014b. InfoWorks CS Help Documentation [Version 14.0]. Innovyze, Oxfordshire, UK.

IPCC, 2007. Climate Change 2007. In: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC). (M.L. Parry, O.F. Canziani, J.P. Palutikof, P.J. van der Linden, and C.E. Hanson, eds.). Cambridge University Press, Cambridge, UK. 1000 p.

IPCC, 2012. Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. (C.B. Field, V. Barros, T.F. Stocker, and Q. Dahe, eds.). Cambridge University Press, Cambridge, UK and New York, USA. 582 p.

ISAACSON, E., STOKER, J.J., TROESCH, A., 1954. Numerical Solution Of Flood Prediction and River Regulation Problems, Report II. New York University Institute of Mathematical Science. Report No. IMM-205, New York, NY.

ISO (INTERNATIONAL STANDARDS ORGANISATION), 1998. ISO 1100/2 1998, Stage-discharge Relation. International Standards Organisation, Geneva.

ITWH, 2000. KOSIM 6.2. Anwenderhandbuch. Institut für technisch-wissenschaftliche Hydrologie GmbH, Hannover, Germany. [în German]

JAIN, A.K., DUBES, R.C., 1988. Algorithms for clustering data. Prentice Hall, Englewood Cliffs, New Jersey.

JAIN, S.K., CHALISGAONKAR, D., 2000. Setting up stage-discharge relations using ANN. Journal of Hydrologic engineering 5, pp. 428-433.

JAKEMAN, A.J., YOUNG, P.C., 1979. Refined instrumental variable methods of recursive time-series analysis, Part II: Multi-variable Systems. International Journal of Control 29, pp. 621-644.

JAKEMAN, A.J., YOUNG, P.C., 1981. Recursive filtering and the inversion of ill-posed causal problems, Utilitas Math 25, pp. 351-376.

JAKEMAN, A.J., YOUNG, P.C., 1984. Recursive filtering and the inversion of ill-posed causal problems. Utilitas Mathematica 35, pp. 351-376.

JAKEMAN, A.J., LITTLEWOOD, I.G., WHITEHEAD, P.G., 1990. Computation of the instantaneous unit hydrograph and identifiable component flows with application to two small upland catchments. Journal of Hydrology 117, pp. 275-300.

JAKEMAN, A.J., HORNBERGER, G.M., 1993. How much complexity is warranted in a rainfall-runoff model? Water Resources Research 29 (8), pp. 2637-2649.

JAKEMAN, A.J., LETCHER, R.A., 2003. Integrated assessment and modelling: features, principles and examples for catchment management. Environmental Modelling & Software 18, pp. 491-501.

JANG, J.-S.R., 1993. ANFIS: adaptive network-based fuzzy inference system. IEEE Transactions on Systems, Man and Cybernetics 23, pp. 665-685.

JANG, J.-S.R., SUN, C.-T., MIZUTANI, E., 2002. Neuro-Fuzzy and Soft Computing. Prentice Hall of India Private Limited, New Delhi, India.

JEKABSONS, G., 2010. M5PrimeLab: M5' Regression Tree and Model Tree Toolbox for Matlab/Octave. Available at: http://www.cs.rtu.lv/jekabsons/, last accessed 27.05.2014.

JONES, B.E., 1916. A method of correcting river discharge for a changing stage, U.S. Geological Survey Water Supply Paper, 375-E, pp. 117-130.

JONKMAN, S.N., BOČKARJOVA, M., KOK, M., BERNARDINI, P., 2008. Integrated hydrodynamic and economic modeling of flood damage in the Netherlands. Ecological Economics 66 (1), pp. 77-90.

KALTEH, A.M., 2013. Monthly river flow forecasting using artificial neural network and support vector regression models coupled with wavelet transform. Computers & Geosciences 54, pp. 1-8.

KANSO, A., GROMAIRE, M.-C., GAUME, E., TASSIN, B., CHEBBO, G., 2003. Bayesian approach for the calibration of models: application to an urban stormwater pollution model. Water Science and Technology 47 (4), pp. 77-84.

KARRA, S., KARIM, M.N., 2009. Comprehensive methodology for detection and diagnosis of oscillatory control loops. Control Engineering Practice 17 (8), pp. 939-956.

KEEFER, T.N., MCQUIVERY, R.S., 1974. Multiple linearization flow routing model. Journal of the Hydraulics Division 100 (HY7), pp. 1031-1046.

KEUPERS, I., WILLEMS, P., 2012. Impact of urban WWTP and CSO fluxes on river peak flow extremes under current and future climate conditions. Water Science and Technology 67 (12), pp. 2670-2676.

KEUPERS, I., WILLEMS, P., 2015a.Global sensitivity analysis of transformation processes in a river water quality model by means of conceptualization. In: E-Proceedings of the 36th IAHR World Congress, 28 June – 3 July 2015, The Hague, the Netherlands.

KEUPERS, I., WILLEMS, P., 2015b. CSO water quality generator based on calibration to WWTP influent data. In: Maere, T., Duchesne, S., Vanrolleghem, P. (Eds.), Proceedings of the 10th Urban Drainage Modelling Conference, 20-23 September 2015, Québec, Canada.

KEUPERS, I., WOLFS, V., KROLL, S., WILLEMS, P. (2015). Impact analysis of sewer overflows on the receiving river water quality using an integrated conceptual model. In: Maere, T., Duchesne, S., Vanrolleghem, P. (Eds.), Proceedings of the 10th Urban Drainage Modelling Conference, 20-23 September 2015, Québec, Canada.

KHAN, M.H., 1993. Muskingum flood routing model for multiple tributaries. Water Resources Research 29, pp. 1057-1062.

KLEIDORFER, M., DELETIC, A., FLETCHER, T.D., RAUCH, W., 2009. Impact of input data uncertainties on urban stormwater model parameters. Water Science and Technology 60, pp. 1545-1554.

KLOPROGGE, P., VAN DER SLUIJS, J., WARDEKKER, J. 2007. Uncertainty communication. Issues and good practice. Copernicus institute for sustainable development and innovation: 60 p.

KOUSSIS, A.D., 1978. Theoretical estimations of flood routing parameters. Journal of the Hydraulic Division 104 (HY1), pp. 109-115.

KOUSSIS, A.D., 1980. Comparison of Muskingum method difference schemes. Journal of the Hydraulic Division 106 (HY5), pp. 925-929.

KOUSSIS, A.D., 2010. Comment on "A praxis-oriented perspective of streamflow inference from stage observations – the method of Dottori et al. (2009) and the alternative of the Jones Formula, with the kinematic wave celerity computed on the looped rating curve by Koussis (2009). Hydrology and Earth System Sciences 14, pp. 1093-1097.

KRON, W., 2002. Flood Risk = Hazard X Exposure X Vulnarability, in Wu et al. (eds.) Flood defence 2002, Science Press, New York Ltd., ISBN 7-03-008310-5.

KRON, W., 2005. Flood Risk = Hazard · Values · Vulnerability. Water International 30 (1), pp. 58-68.

KRUEGER, T., PAGE, T., HUBACEK, K., SMITH, L., HISCOCK, K., 2012. The role of expert opinion in environmental modelling. Environmental Modelling and Software 36, pp. 4-18.

KRZYSZTOFOWICZ, R. 1999. Bayesian theory of probabilistic forecasting via deterministic hydrologic model. Water Resources Research 35, pp. 2739-2750.

KRZYSZTOFOWICZ, R., 2002. Bayesian system for probabilistic river stage forecasting. Journal of Hydrology 268, pp. 16-40.

KULANDAISWAMY, V.C., 1964. A basic study of the rainfall excess-surface runoff relationship in a basin system. Ph.D. Thesis, University of Illinois, Urbana.

KUMAR, D.N., BALIARSINGH, F., RAJU, K.S., 2011. Extended Muskingum method for flood routing. Journal of Hydro-environment Research 5, pp. 127-135.

KUO, B.C., GOLNARAGHI, F., 2002. Automatic Control Systems. New York, NY, USA: John Wiley & Sons, Inc.

LAMMERSEN, R., ENGEL, H., VAN DE LANGEMHEEN, W., BUITEVELD, H., 2002. Impact of river training and retention measures on flood peaks along the Rhine. Journal of Hydrology 267, pp. 115-124.

LANIAK, G.F., OLCHIN, G., GOODALL, J., VOINOV, A., HILL, M., GLYNN, P., WHELAN, G., GELLER, G., QUINN, N., BLIND, M., PECKHAM, S., REANEY, S., GABER, N., KENNEDY, R., HUGHES, A., 2013. Integrated environmental modeling: A vision and roadmap to the future. Environmental Modelling & Software 39, pp. 3-23.

LAWSON, E., THORNE, C., AHILAN, S., ALLEN, D., ARTHUR, S., EVERETT, G., FENNER, R., GLENIS, V., GUAN, D., HOANG, L., KILSBY, C., LAMOND, J., MANT, J., MASKREY, S., MOUNT, N., SLEIGH, A., SMITH, L., WRIGHT, N., 2014. Delivering and evaluating the multiple flood risk benefits in Blue-Green Cities: an interdisciplinary approach. In: Flood Recovery, Innovation and Response IV (D. Proverbs, C.A. Brebbia, eds.), pp. 113-124. WIT press.

LEES, M.J., 2000. Data-based mechanistic modelling and forecasting of hydrological systems. Journal of Hydroinformatics 2 (1), pp. 15-34.

LI, X., WANG, J., HUANG, B., LU, S., 2010. The DCT-based oscillation detection method for a single time series. Journal of Process Control 20, pp. 609-617.

LIGGETT, J.A., CUNGE, J.A., 1975.Numerical methods of solutions of the unsteady flow equations, Unsteady Flow in Open Channels. Water Resources Publications, Fort Collins, Chapter 4.

LIGHTHILL, M.J., WHITHAM, G.B., 1955. On kinematic floods, I: Flood movements in long rivers. Proceedings of Royal Society London, A229, pp. 281-316.

LINSLEY, R.K., KOHLER, M.A., PAULHUS, J.L.H., 1949. Applied Hydrology. New York: McGraw Hill Book Company, pp. 502-530.

LOHANI, A.K., GOEL, N.K., BHATIA, K.K.S, 2006. Takagi-Sugeno fuzzy inference system for modeling stage-discharge relationship. Journal of Hydrology 331, pp. 146-160.

MACKAY, D.J.C., 1992. Bayesian Interpolation. Neural Computation 4, pp. 415-447.

MADSEN, H. 2000. Automatic calibration of a conceptual rainfall-runoff model using multiple objectives. Journal of Hydrology 235, pp. 276-288.

MAIER, H.R., DANDY, G.C., 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environmental Modelling & Software 15 (1), pp. 101-124.

MAIER, H.R., JAIN, A., DANDY, G., SUDHEER, K.P., 2010. Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions. Environmental Modelling & Software 25, pp. 891-909.

MAKROPOULOS, C.K., NATSIS, K., LIU, S., MITTAS, K., BUTLER, D., 2008. Decision support for sustainable option selection in integrated urban water management. Environmental Modelling & Software 23, pp. 1448-1460.

MAMDANI, E.H., 1977. Application of fuzzy logic to approximate reasoning using linguistic systems. Fuzzy Sets and Systems 26, pp. 1182-1191.

MANDER, R.J., 1978. Aspects of unsteady flow and variable backwater. In: Hydrometry: Principles and Practice (first edn). Ed. Herschy. John Wiley & Sons, Chichester, UK, 511 p.

MANNINA, G., VIVIANI, G., FRENI, G., 2004. Modelling the integrated urban drainage systems. In: Bertrand-Krajewski, J.-L., Almeida, M., Matos, J., Abdul-Talib, S. (Eds.), Sewer Networks and Processes within Urban Water Systems. IWA Publishing, London, pp. 3–12.

MCCARTHY, G.T., 1938. The unit hydrograph and flood routing. Paper presented at the Conferense of the North Atlantic Division of US Corps of Engineering, New London, Connecticut.

MCLEAN, R. F. , TSYBAN,A., BURKETT ,A., CODIGNOTTO, J.O., FORBES, D.L., MIMURA, N., BEAMISH, R.J., ITTEKKOT, V., 2001. Coastal zones and marine

ecosystems. In:McCarthy, J.J.,Canziani, O.F., Leary, N.A., Dokken, D.J., and White, K.S. (eds.), Climate Change 2001: Impacts, Adaptation and Vulnerability. Cambridge: Cambridge University Press, pp. 343–380.

MEERT, P., PEREIRA, F., WILLEMS, P., 2016. Computationally efficient modeling of tidal rivers using conceptual reservoir-type models. Environmental Modelling & Software 77, pp. 19-31.

MEIRLAEN, J., HUYGHEBAERT, B., SFORZI, F., BENEDETTI, L., VANROLLEGHEM, P., 2001. Fast, simultaneous simulation of the integrated urban wastewater system using mechanistic surrogate models. Water Science and Technology 47 (7), pp. 301-308.

MEUNIER, S., LITRICO, X., BELAUD, G., 2007. Linear approximation of open-channel flow routing with backwater effect. 32nd Congress on the International Association of Hydraulic Engineering and Research 2007, Venice, Italy, 10 p.

MIAO, T., SEBORG, D.E., 1999. Automatic detection of excessively oscillatory feedback control loops. In: Proceedings of the 1999 IEEE international conference on control applications, HI, USA, pp. 359-364.

MILLER, J.E., 1984. Basic concepts of kinematic-wave models. Professional Paper 1302. USGS Numbered Series, 29 p.

MILLER, W.A., CUNGE, J.A., 1975. Simplified equations of unsteady flow. In Mahmood, K., and Yevjevich, V. (eds). Unsteady Flow in Open Channels. Volume I , Fort Collins, Colorado: Water Resources Publications, Chapter 5, pp. 183-257.

MILLER, W.A., YEVJEVICH, V., 1975. Unsteady flow in Open Channels. Bibliography, Vol. III. Fort Collins, Colorado: Water Resources Publications.

MIRA, 2011. Milieurapport Vlaanderen, Achtergronddocument 2010, Waterkwantiteit. Peeters, B., Defloor, W., De Jongh, I., D'hont, D., Fronhoffs, A., Lermytte, J., Ottoy, N., Vanhille, A., Verlé, W., Voet, M., Michielsen, S., Vanneuville, W., Vlaamse Milieumaatschappij. Raadpleegbaar op www.milieurapport.be. [in Dutch]

MIRA, 2015. Klimaatrapport 2015: over waargenomen en toekomstige klimaatveranderingen. Brouwers, J., Peeters, B., Van Steertegem, M., Van Lipzig, N., Wouters, H., Beullens, J., Demuzere, M., Willems, P., De Ridder, K., Maiheu,

B., De Troch, R., Termonia, P., Vansteenkiste, T., Craninx, M., Maetens, W., Defloor, W., Cauwenberghs, K., Vlaamse Milieumaatschappij. Raadpleegbaar op www.milieurapport.be. [in Dutch]

MITCHELL, V.G., 2004. Integrated Urban Water Management, a Review of Current Australian Practice. Urban Water Systems & Technologies. CSIRO.

MITCHELL, V.G., MEIN, R.G., MCMAHON, T.A., 2001. Modelling the urban wwater cycle. Environmental Modelling & Software 16, pp. 615-629.

MOLINARO, P., FILIPPO, A.D., FERRARI, F., 1994. Modelling of flood wave propagation over flat areas of complex topography in presence of different infrastructures. In: Proceedings of the Speciality Conference on Modelling of Flood Propagation over Initially Dry Areas, (Molinaro, P., Natale, L., Ed.), ASCE, New York, NY, USA, Milan, Italy, pp. 209-228.

MONTANARI, A. & GROSSI, G., 2008. Estimating the uncertainty of hydrological forecasts: A statistical approach. Water Resources Research 44, W00B08.

MONTAZAR, A., RIAZI, H., BEHBAHANI, S.M., 2010. Conjunctive Water Use Planning in an Irrigation Command Area. Water Resources Management 24 (3), pp. 577-596.

MOORE, R.J., 2007 The PDM rainfall-runoff model. Hydrology & Earth System Sciences 11 (1), pp. 483-499.

MOTIEE, H., CHOCAT, B., BLANPAIN, O., 1997. A storage model for the simulation of the hydraulic behavior of drainage networks. Water Science and Technology 36 (8-9), pp. 57-63.

MOUSSA, R., 1996. Analytical Hayami solution for the diffusive wave flood routing problem with lateral inflow. Hydrological Processes 10, pp. 1209-1227.

MUSCHALLA, D., REUSSNER, F., SCHNEIDER, S., OSTROWSKI, M.W., 2006. Dokumentation des Schmutzfrachtmodells SMUSI Version 5.0. Institut für Wasserbau und Wasserwirtschaft, Technische Unversität Darmstadt, Germany. [in German].

MUSCHALLA, D., HOFER, T., MAIER, R., RIEGER, L., SCHRAA, O., GRUBER, G., 2015. Development of an innovative concept for building integrated urban water

models. In: Maere, T., Tik, S., Duchesne, S., Proceedings of the 10th Urban Drainage Modelling Conference, 20-23 September 2015, Québec, Canada.

NASH, J.E., 1957. The form of the instantaneous unit hydrograph. In: General Assembly of Toronto, Vol. III Surface Waters, Prevision, Evaporation, pp. 114-121: IAHS Publication No. 45.

NASH, J.E., 1959. A note on the Musking method of flood routing. Journal of Geophysical Research 64, pp. 1053-1056.

NASH, J.E., SUTCLIFFE, J.V., 1970. River flow forecasting through conceptual models part I – A discussion of principles. Journal of Hydrology 10 (3), pp. 282-290.

NGUYEN, T.G., DE KOK, J.L., TITUS, M.J., 2007. A new approach to testing an integrated water systems model using qualitative scenarios. Environmental Modelling and Software 22 (11), pp. 1557-1571.

NORTON, J.P., 1986. An Introduction to Identification. Dover Books on Electrical Engineering, Mineola, NJ, USA. 320 p.

PANSIC, N., YEN, B. C., 1982. Kinetic-wave modeling of storm sewers with surcharge. Urban stormwater quality, management and planning. Proceedings of the second international conference on urban storm drainage. 193-199. Littleton, CO: Water Resources Publications.

PARKER, P., LETCHER, R., JAKEMAN, A., BECK, M.B., HARRIS, G., ARGENT, R.M., HARE, M., PAHL-WOSTL, C., VOINOV, A., JANSSEN, M., SULLIVAN, P., SCOCCIMARRO, M., FRIEND, A., SONNENSHEIN, M., BARKER, D., MATEJICEK, L., ODULAJA, D., DEADMAN, P., LIM, K., LAROCQUE, G., TARIKHI, P., FLETCHER, C., PUT, A., MAXWELL, T., CHARLES, A., BREEZE, H., NAKATANI, N., MUDGAL, S., NAITO, W., OSIDELE, O., ERIKSSON, I., KAUTSKY, U., KAUTSKY, E., NAESLUND, B., KUMBLAD, L., PARK, R., MALTAGLIATI, S., GIRARDIN, P., RIZZOLI, A., MAURIELLO, D., HOCH, R., PELLETIER, D., REILLY, J., OLAFSDOTTIR, R., BIN, S., 2002. Progress in integrated assessment and modelling. Environmental Modelling & Software 17, pp. 209-217.

PARMET, B., BUISHAND, T., BRANDSMA, T., MÜLDERS, R., 1999. Design discharge of the large rivers in The Netherlands – towards a new methodology.

Hydrological Extremes: Understanding, Predicting, Mitigating. Proceedings of IUGG 99 Symposium HIS, Birmingham, July 1999. IAHS Publ. no. 255.

PEDREGAL, D.J., RIVAS, R., FELIU, V., SÁNCHEZ, L., LINARES, A., 2009. A non-linear forecasting system for the Ebro River at Zaragoza, Spain. Environmental Modelling and Software 24 (4), pp. 502-509.

PERALTA, R., CANTILLER, R., TERRY, J.E., 1995. Optimal Large-Scale Conjunctive Water-Use Planning: Case Study. Journal of Water Resources Planning and Management 121 (6), pp. 471-478.

PERUMAL, M., SHRESTHA, K.B., CHAUBE, U.C., 2004. Reproduction of Hysteresis in Rating Curves, Journal of Hydrologic Engineering – ASCE 130, 870-878.

PETERSEN-ØVERLEIR, A., 2006. Modelling stage-discharge relationships affected by hysteresis using the Jones formula and nonlinear regression. Hydrological Sciences Journal 51 (3), pp. 365-388.

PIEGAT, A., 2001. Fuzzy Modeling and Control. Physica-Verlag, Heidelberg, New York.

PIETZ, D., 2002. Engineering the State: The Huai River and Reconstruction in Nationalist China 1927-1937. Routledge. ISBN 0-415-93388-9. pg. xvii, pg. 61-70.

POELMANS, L., VAN ROMPAEY, A., BATELAAN, O., 2010. Coupling urban expansion models and hydrological models: How important are spatial patterns? Land Use Policy 27, pp. 965-975.

PONCE, V.M., YEVJEVICH, V., 1978. Muskingum-Cunge method with variable parameters. Journal of the Hydraulic Division 104 (HY12), pp. 1663-1667.

PONCE, V.M., CHAGANTI, P.V., 1994. Variable-parameter Muskingum-Cunge method revisited. Journal of Hydrology, 162, pp. 433-439.

PORPORATO, A., RIDOLFI, L., 2001. Multivariate nonlinear prediction of river flows. Journal of Hydrology 248 (1-4), pp. 109-122.

PULS, L.G., 1928. Construction of Flood Routing Curves. House Document 185, US 70th Congress, 1st session, Washington DC, pp. 46-52.

QUICK, M.C., PIPES, A., 1975. Nonlinear channel routing by computer. Journal of the Hydraulic Division 101 (HY6), pp. 651-664.

QUINLAN, J.R., 1992. Learning with continuous classes. In: Proceedings of the 5th Australian Joint Conference on Artificial Intelligence. World Scientific, Singapore, pp. 343-348.

RAFTERY, A.E. 1993. Bayesian model selection in structural equation models. In: Bollen, K.A. & Long, J.S. (ed.), Testing Structural Equation Models, Sage, Beverly Hills, CA., pp. 163-180.

RAUCH, W., HARREMOËS, P., 1996. The importance of the treatment plant performance during rain to acute water pollution. Water Science and Technology 34 (3-4), pp. 1-8.

RAUCH, W., BERTRAND-KRAJEWSKI, J.-L., KREBS, P., MARK, O., SCHILLING, W., SCHÜTZE, M., VANROLLEGHEM, P.A., 2002. Deterministic modelling of integrated urban drainage systems. Water Science and Technology 45 (3), pp. 81-94.

RAVAZI, S., TOLSON, B.A., BURN, D.H., 2012. Review of surrogate modeling in water resources. Water Resources Research 48 (7). W07401.

REFSGAARD, J.C., MADSEN, H., ANDREASSIAN, V., ARNBJERG-NIELSEN, K., DAVIDSON, T.A., DREWS, M., HAMILTON, D.P., JEPPESEN, E., KJELLSTROM, E., OLESEN, J.E., SONNENBORG, T.O., TROLLE, D., WILLEMS, P., CHRISTENSEN, J.H., 2014. A framework for testing the ability of models to project climate change and its impacts. Climatic Change 122 (1-2), pp. 271-282.

REICHLE, R.H., MCLAUGHLIN, D.B., ENTEKHABI, D., 2002. Hydrological Data Assimilation with the Ensemble Kalman Filter. Montrly Weather Review 130, pp. 103-114.

REJANI, R., JHA, M.K., PANDA, S.N., 2009. Simulation-optimization modelling for sustainable groundwater management in a coastal basin of Orissa, India. Water Resources Management 23, pp. 235-263.

ROMANOWICZ, R.J., YOUNG, P.C., BEVEN, K.J., PAPPENBERGER, F., 2008. A data based mechanistic approach to nonlinear flood routing and adaptive flood level forecasting. Advances in Water Resources 31, pp. 1048-1056.

RÖSSERT, R., 1964. Hydraulik in Wasserbau, Oldenbourg R. [in German]

SAINT-VENANT, B. DE, 1871. Theory of unsteady water flow, with application to river floods and to propagation of tides in river channels. Computes Rendus Acad. Sci., Paris, 73, 148-154, 237-240. (Translated into English by US Corps of Engineers, No. 49-g, Waterways Experiment Station, Vicksburg, Mississippi, 1949).

SALSBURY, T.I., SINGHAL, A., 2005. A new approach for ARMA pole estimation using higher-order crossings. In: Proceedings of the 2005 American Control Conference, Portland, OR, USA, pp. 4458-4463.

SALVADORE, E., BRONDERS, J., BATELAAN, O., 2015. Hydrological modelling of urbanized catchments: a review and future directions. Journal of Hydrology 529 (1), pp. 62-81.

SARTOR, J., 1999. Simulating the influence of backwater effects in sewer systems using hydrological model components. Water Science and Technology 39 (9), pp. 145-152.

SARVI, M., AVANAKI, I.N., 2015. An optimized Fuzzy Logic Controller by Water Cycle Algorithm for power management of Stand-alone Hybrid Green Power generation. Energy Conversion and Management 106, pp. 118-126.

SAVCENCO, V., HUNDSDORFER, W., VERWER, G., 2007. A multirate time stepping strategy for stiff ordinary differential equations. BIT Numerical Mathematics 47, pp. 137-155.

SCHLEGEL, M., KNOTH, O., ARNOLD, M., WOLKE, R., 2009. Multirate Runge-Kutta schemes for advection equations. Journal of Computational and Applied Mathematics 226, pp. 345-357.

SCHÜTZE, M., BUTLER, D., BECK, B., 1996. Development of a framework for the optimisation of runoff, treatment and receiving waters. In: Proceedings 7th International Conference on Urban Storm Drainage. Hannover, Germany, September 9-13 1996. pp. 1419-1425.

SCHÜTZE, M., CAMPISANO, A., COLAS, H., SCHILLING, W., VANROLLEGHEM, P.A., 2004. Real time control of urban wastewater systems – where do we stand today? Journal of Hydrology 299, pp. 335-348.

SHAMSELDIN, A.Y., 1997. Application of a neural network technique to rainfall-runoff modelling. Journal of Hydrology 199, pp. 272-294.

SHAW, E.M., 1983. Hydrology in Practice. Capman and Hall, London, UK. 569 p.

SIMONS, D.B., RICHARDSON, E.V., 1962. The effect of bed roughness on depth-discharge relations in alluvial channels. U.S. Geological Survey Water-Supply paper 1498-E.

SINGH, K.K., PAL, M., SINGH, V.P., 2010. Estimation of mean annual flood in Indian catchments using back propagation neural network and M5 model tree. Water Resources Management 24, pp. 2007-2019.

SMALL, S.J., JAY, L.O., MANTILLA, R., CURTU, R., CUNHA, L.K., FONLEY, M., KRAJEWSKI, W.F., 2013. An asynchronous solver for systems of ODEs linked by directed tree structure. Advances in Water Resources 53, pp. 23-32.

SOLOMATINE, D.P., DULAL, K.N., 2003. Model tree as an alternative to neural network in rainfall-runoff modelling. Hydrological Sciences Journal 48 (3), pp. 399-411.

SOLOMATINE, D.P., YUNPENG, X., 2004. M5 model trees and neural networks: application to flood forecasting in the upper reach of the Huai River in China. Journal of Hydrologic Engineering 9 (6), pp. 491-501.

SOLVI, A.-M., 2007. Modelling the Sewer-Treatment-Urban River System in view of the EU Water Framework Directive. PhD thesis, Ghent University, Belgium, pp. 218.

SOROOSHIAN, S., GUPTA, K.V., 1995. Model Calibration. In: Computer models of watershed hydrology. Chapter 2. Singh, V.P. (Ed.). Water Resources Publications.

SPOLIA, S.K., CHANDER, S., 1974. Modelling of surface runoff systems by an ARMA model. Journal of Hydrology 22, pp. 317-332.

STEWART, M.D., BATES, P.D., ANDERSON, M.G., PRICE, D.A., BURT, T.P., 1999. Modeling floods in hydrologically complex lowland river reaches. Journal of Hydrology 223 (1-2), pp. 85-106.

STOKER, J.J., 1953. Numerical Solution of Flood Prediction and River Regulation Problems; Derivation of Basic Theory and Formulation of Numerical Methods of Attack. Report I, New York University Institute of Mathematical Science, Report No. IMM-NYU-200, New York, NY.

STRUPCZEWSKI, W., KUNDZEWICZ, Z., 1980. Translatory characteristics of the Musking method of flood routing – a comment. Journal of Hydrology 98, pp. 363-368.

STURM, T.W., 2001. Open Channel Hydraulics. McGraw-Hill.

SUGENO, M., KANG, G.T., 1988. Structure identification of fuzzy model. Fuzzy Sets and Systems 28, pp. 15-33.

SWAMEE, P.K., 1992. Sluice-gate discharge equations. Journal of Irrigation and Drainage Engineering 118 (1), pp. 56-60.

SWENNEN, B., VREYS, C., 2010. Adaptive measures against floods in the Dender catchment. MSc. Thesis, KU Leuven, Faculty of Engineering, Belgium. [in Dutch]

TABARI, H., TAYE, M.T., WILLEMS, P., 2015. Actualisatie en verfijning klimaatscenario's tot 2100 voor Vlaanderen – Appendix 2: Nieuwe modelprojecties voor Ukkel op basis van globale klimaatmodellen (CMIP5). Studie uitgevoerd in opdracht van de Afdeling Operationeel Waterbeheer van de Vlaamse Milieumaatschappij en MIRA, MIRA/2015/03, KU Leuven. Raadpleegbaar op www.milieurapport.be.

TAKAGI, T., SUGENO, M., 1985. Fuzzy identification of systems and its application to modeling and control. IEEE Transactions on Systems, Man and Cybernetics 15 (1), pp. 116-132.

TATUM, F.E., 1940. A Simplified Method of Routing Flood Flows Through Natural Valley Storage. Unpublished memo, US Engineers Office, Rock Island, Illinois.

TAWFIK, M., IBRAHIM, A., FAHMY, H., 1997. Hysteresis Sensitive Neural Network for Modeling Rating Curves. Journal of Computing in Civil Engineering 11, pp. 206-211.

TAYLOR, C.J., PEDREGAL, D.J., YOUNG, P.C., TYCH, W., 2007. Environmental time series analysis and forecasting with the Captain toolbox. Environmental Modelling and Software 22 (6), pp. 797-814.

THIELE, B., OTTER, M., MATTSON, S.E., 2014. Modular Multi-Rate and Multi-Method Real-Time Simulation. In: Proceedings of the 10th International Modelica Conference, Lund, Sweden, March 10-12, 2014.

THORNHILL, N.F., HÄGGLUND, T., 1997. Detection and diagnosis of oscillations in control loops. Control Engineering Practice 5 (10), pp. 1343-1354.

TIMBE, L., 2007. River flooding analysis using quasi-2D hydraulic modelling and geospatial data. PhD thesis, KU Leuven, Faculty of Engineering, Belgium.

TODINI, E., 2008. A model conditional processor to assess predictive uncertainty in flood forecasting. International Journal of River Basin Management 6 (2), pp. 123-137.

TODINI, E., WALLIS, J.R., 1977. Using CLS for Daily or Longer Period Rainfall-Runoff Modelling. Mathematical Models for Surface Water Hydrology, edited by Ciriani, T.A., Maione, U. and Wallis, J.R., John Wiley & Sons, London. pp. 149-168.

TSAI, C.W., 2003. Applicability of kinematic, noninertia, and quasi-steady dynamic wave models to unsteady flow routing. Journal of Hydraulic Engineering, 129 (8), pp. 613-627.

TSAI, C.W., 2005. Flood routing in mild-sloped rivers – wave characteristics and downstream backwater effect. Journal of Hydrology 301 (1-4), pp. 151-167.

URICH, C., RAUCH, W., 2014. Modelling the urban water cycle as an integrated part of the city: a review. Water Science and Technology 70 (11), pp. 1857-1872.

VAES, G., 1999. The influence of rainfall and model simplification on combined sewer system design. PhD Thesis, Department of Civil Engineering. University of Leuven, Belgium.

VAES, G., BERLAMONT, J., 1996. Composite storms as rainfall input for sewer calculations. Tijdschrift Water 88, pp. 143-148. [în Dutch]

VAES, G., WILLEMS, P., BERLAMONT, J., 2000. Selection and composition of representative hydrographs for river design calculations. International Conference on "Monitoring and modelling catchment water quantity and quality" (ERB 2000), Gent, 27-29 Sept. 2000, pp. 53-55.

VAN DEN ZEGEL, B., VERMUYTEN, E., WOLFS, V., MEERT, P., WILLEMS, P., 2014. Real-time control of floods along the Demer River, Belgium, by means of MPC in combination with GA and a fast conceptual river model. In: Proceedings of the 11th International Conference on Hydroinformatics, HIC2014. New York City, USA, August 17-21 2014.

VAN HOOF, K., 2003. Het SENTWA-model: een berekening van de nutriëntbelasting van oppervlaktewater. Water 11-12, pp. 1-6. [în Dutch]

VAN STEENBERGEN, N., VERBELEN, J., 2009. Adaptive measures against floods in the Dender catchment. MSc. Thesis, KU Leuven, Faculty of Engineering, Belgium.

VAN STEENBERGEN, N., RONSYN, J., WILLEMS, P., VAN EERDENBRUGH, K., 2011. A data-based probabilistic approach to calculate and visualize the uncertainty of flood forecasts. In Zenz, G. (ed.), Hornich, R. (ed.), Urban Flood Risk Management (UFRIM 2011) 'Approaches to enhance resilience of communities'. Graz, Austria, 21-23 September 2011. pp. 177-182.

VAN STEENBERGEN, N., RONSYN, J., WILLEMS, P., 2012. A non-parametric data-based approach for probabilistic flood forecasting in support of uncertainty communication. Environmental Modelling & Software 33, pp. 92-105.

VAN STEENBERGEN, N., WILLEMS, P., 2014. Rainfall uncertainty in flood forecasting: Belgian case study of Rivierbeek. Journal of Hydrologic Engineering 19 (10), 05014013.

VANROLLEGHEM, P.A., FRONTEAU, C., BAUWENS, W., 1996. Evaluation of design and operation of the sewage transport and treatment system by an EQO/EQS based analysis of the receiving water immission characteristics. In: Proceedings WEF Conference on Urban Wet Weather Pollution. Québec, Canada, June 16-19 1996. pp. 14.35-14.46.

VANROLLEGHEM, P.A., SCHILLING, W., RAUCH, W., KREBS, K., AALDERINK, H., 1999. Setting up measuring campaigns for integrated wastewater modelling. Water Science and Technology 39 (4), pp. 257-268.

VANROLLEGHEM, P.A., BENEDETTI, L., MEIRLAEN, J., 2005. Modelling and real-time control of the integrated urban wastewater system. Environmental Modelling & Software 20, pp. 427-442.

VANROLLEGHEM, P.A., KAMRADT, B., SOLVI, A.-M., MUSCHALLA, D., 2009. Making the best of two hydrological flow routing models: Nonlinear outflow-volume relationships and backwater effects model. In: Proceedings 8th International Conference on Urban Drainage Modelling (8UDM). Tokyo, Japan, September 7-11 2009.

VANSTEENKISTE, T., PEREIRA, F., WILLEMS, P., VANNEUVILLE, W., VAN EERDENBRUGH, K., MOSTAERT, F., 2011a. Effect of climate change on the hydrological regime of navigable water courses in Belgium. Subreport I – Literature review on climate research, 43 p. Antwerpen, Belgium: Flanders Hydraulics Research.

VANSTEENKISTE, T., PEREIRA, F., WILLEMS, P., MOSTAERT, F., 2011b. Effect of climate change on the hydrological regime of navigable water courses in Belgium. Subreport II – Climate change impact analysis by conceptual models. 57 p. Antwerpen, Belgium: Flanders Hydraulics Research.

VANSTEENKISTE, T., TAVAKOLI, T., NTEGEKA, V., DE SMEDT, F., BATELAAN, O., PEREIRA, F., WILLEMS, P., 2014. Intercomparison of hydrological model structures and calibration approaches in climate scenario impact projections. Journal of Hydrology 519, pp. 743-755.

VERMUYTEN, E., VAN DEN ZEGEL, B., WOLFS, V., MEERT, P., WILLEMS, P., 2014. Real-time flood control by means of an improved MPC-GA algorithm and a fast conceptual river model for the Demer basin in Belgium. Proceedings of the 6th International Conference on Flood Management (ICFM). Sao Paulo, Brasil, September 2014, pp. 1-8.

VERNIEUWE, H., DE BAETS, B., VERHOEST, N., 2006. Comparison of clustering algorithms in the identification of Takagi-Sugeno models: A hyrdological case study. Fuzzy Sets and Systems 157, pp. 2876-2896.

VEZZARO, L., CHRISTENSEN, M.L., THIRSING, C., GRUM, M., MIKKELSEN, P.S., 2014. Water quality-based real time control of integrated urban drainage systems: a preliminary study from Copenhagen, Denmark. Procedia Engineering 70, pp. 1707-1716.

VIESSMAN, W., LEWIS, G., KNAPP, J.W., 1989. Introduction to hydrology. Harper & Row Publishers, 3rd edn, New York, USA, 780 p.

VILLAZON, M., 2011. Modelling and conceptualization of hydrology and river hydraulics in flood conditions, for Belgian and Bolivian basins. PhD thesis, KU Leuven, Faculty of Engineering, Belgium.

VISSERS, S., WOLFS, V., 2011. Adaptive measures against floods in the Dender catchment. MSc. Thesis, KU Leuven, Faculty of Engineering, Belgium. [în Dutch]

VLACHOS, E., BRAGA, B., (EDS.), 2001. The Challenge of Urban Water Management. IWA Publishing, Cornwall, UK.

VLAREM, 2010. VLAREM II: Legislation Environment, Nature and Energy (Bijlage 2.3.1. Basismilieukwaliteitsnormen voor oppervlaktewater). Flemish Government of Belgium. [în Dutch]

VMM, 2006. Grondwaterbeheer in Vlaanderen: het onzichtbare water doorgrond. Vlaamse Milieumaatschappij, Aalst. 150 p. [in Dutch]

VOINOV, A., SHUGART, H., 2013. 'Integronsters', integral and integrated modeling. Environmental Modelling and Software 39, pp. 149-158.

VRUGT, J.A., GUPTA, H.V., BOUTEN, W., SOROOSHIAN, S., 2003a. A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. Water Resources Research 39 (8), 1201.

VRUGT, J.A., GUPTA, H.V., BASTIDAS, L.A., BOUTEN, W., SOROOSHIAN, S., 2003b. Effective and efficient algorithm for multiobjective optimization of hydrologic models. Water Resources Research 38 (8), 1214.

WAGENER, T., WHEATER, H.S., GUPTA, H.V., 2004. Rainfall-Runoff Modeling in Gauged and Ungauged Catchments. Imperial College Press, London, UK.

WAGENINGEN UNIVERSITY & RESEARCH CENTRE (WUR), & MILIEU- EN NATUURPLANBUREAU (MNP), 2007. Eururalis 2.0 CD-rom. Wageningen, The Netherlands: Alterra Wageningen UR.

WALTZ, R.A., MORALES, J.L., NOCEDAL, J., ORBAN, D., 2006. An interior algorithm for nonlinear optimization that combines line search and trust region steps. Mathematical Programming 107 (3), pp. 391-408.

WANG, Y., WITTEN, I.H., 1997. Induction of model trees for predicting continuous classes. In: Proceedings of the Poster Papers of the European Conference on Machine Learning. University of Economics, Faculty of Informatics and Statistics, Prague, Czech Republic.

WEERTS, A.H., WINSEMIUS, H.C., VERKADE, J.S. 2011. Estimation of predictive hydrological uncertainty using quantile regression: examples from the national flood forecasting system (England and Wales), Hydrology and Earth System Sciences 15, pp. 255-265.

WEINMANN, P.E., LAURENSON, E.M., 1979. Approximate flood routing methods: a review. Journal of the Hydraulic Division 105 (HY13), Proceeding Paper 15057.

WILLEMS E., DIERCKX J., GIJSBERS B., WILLEMS P., TIMMERMAN, A., 2005. New calibration and validation study for the SENTWA model. Bodemkundige Dienst van België and KU Leuven – Hydraulics Section, Report for the Flemish Environment Agency (VMM) – Division Water Quality Management. [in Dutch]

WILLEMS, P., 2000. Probabilistic immission modeling of receiving surface waters. PhD thesis, KU Leuven, Faculty of Engineering, Belgium.

WILLEMS, P., 2006. Random number generator or sewer water quality model? Water Science and Technology 54 (6-7), pp. 387-394.

WILLEMS, P., 2008.Quantification and relative comparison of different types of uncertainties in sewer water quality modeling. Water Research 42, pp. 3539-3551.

WILLEMS, P., 2009. A time series tool to support the multi-criteria performance evaluation of rainfall-runoff models. Environmental Modelling and Software 24 (3), pp. 311-321.

WILLEMS, P., 2013. River modelling. Acco Leuven & Den Haag, 268 p., ISBN 978-90-334-9296-9, D/2013/0543/132, NUR 955. [in Dutch]

WILLEMS, P., 2014. Parsimonious rainfall-runoff model construction supported by time series processing and validation of hydrological extremes – Part 1: Step-wise model-structure identification and calibration approach. Journal of Hydrology 510, pp. 578-590.

WILLEMS, P., BERLAMONT, J., 2002. Probabilistic emission and immission modelling: case-study of the combined sewer-WWTP-receiving water system at Dessel (Belgium). Water Science and Technology 45 (3), pp. 117-124.

WILLEMS, P., VAES, G., POPA, C., TIMBE CASTRO, L., BERLAMONT, J., 2002. Quasi 2D river flood modeling. River Flow 2002, International Conference on Fluvial Hydraulics. Louvain-La-Neuve, België, 4-6 Sep 2002, pp. 1253-1259.

WILLEMS, P., BAGUIS, P., NTEGEKA, V., ROULIN, E., 2010. Climate change impact on hydrological extremes along rivers and urban drainage systems in Belgium "CCI-HYDR" final report. Brussels: Belgian Science Policy 2010. 110 p. (Research Program Science for a Sustainable Development).

WILLEMS, P., OLSSON, J., ARNBJERG-NIELSEN, K., BEECHAM, S., PATHIRANA, A., BÜLOW GREGERSEN, I., MADSEN, H., NGUYEN, V-T-V., 2012. Impacts of climate change on rainfall extremes and urban drainage. IWA Publishing, 252p., Paperback Print ISBN 9781780401256; Ebook ISBN 9781780401263.

WOLFS, V., WILLEMS, P., 2013. A data driven approach using Takagi-Sugeno models for computationally efficient lumped floodplain modelling. Journal of Hydrology 503, pp. 222-232.

WOLFS, V., VILLAZON, M., WILLEMS, P., 2013. Development of a semi-automated model identification and calibration tool for conceptual modelling of sewer systems. Water Science and Technology 68 (91), pp. 167-175.

WOLFS, V., WILLEMS, P., 2014. Computationally efficient lumped floodplain modelling using data-driven methods. In: Proceedings of the 11th International Conference on Hydroinformatics, HIC2014. New York City, USA, August 17-21 2014.

WONG, T., BROWN, R., 2009. The water sensitive city: principles for practice. Water Science and Technology 60 (3), pp. 673-682.

WOOLHISER, D.A., 1973. Hydrologic and Watershed Modelling – State of the Art. Transactions of the ASEA. ASAE Paper No. 71-747. doi: 10.13031/2013.37568

WU, C.L., CHAU, K.W., 2011. Rainfall-runoff modelling using artificial neural network coupled with singular spectrum analysis. Journal of Hydrology 399 (3-4), pp. 394-409.

XIE, X., ZHANG, D., 2010. Data assimilation for distributed hydrological catchment modeling via ensemble Kalman filter. Advances in Water Resources 33 (6), pp. 678-690.

YAGER, R.R., FILEV, D.P., 1994. Generation of fuzzy rules by mountain clustering. IEEE Transactions on Systems, Man and Cybernetics 24, pp. 209-219.

YANG, Z., HAN, D., 2006. Derivation of unit hydrograph using a transfer function approach. Water Resources Research 42: W01501.

YAZDI, J., SALEHI NEYSHABOURI, S.A.A., 2014. Adaptive surrogate modeling for optimization of flood control detention dams. Environmental Modelling & Software 61, pp. 106-120.

YEH, W.W.G., 1985. Reservoir management and operation models: a state-of-the-art review. Water Resources Research 21, pp. 1797-1818.

YEN, B.C., 1994. Is hydraulics over-used or under-used in storm drainage. Water Science and Technology 29 (1-2), pp. 53-61.

YEN, J.F., LIN, C.H., TSAI, C.T., 2001. Hydraulic characteristics and discharge control of sluice gates. Journal of the Chinese Institute of Engineers 24 (3), pp. 301-310.

YOUNG, P.C., 1984. Recursive estimation and time series analysis. Berlin: Springer.

YOUNG, P.C., 1989. Recursive estimation, forecasting and adaptive control. In: C.T. Leondes (Ed.), Control and Dynamic Systems. Academic Press, San Diego, CA, pp. 119-166.

YOUNG, P.C., 1993. Time variable and state dependent modelling of nonstationary and nonlinear time series. Chapter 26 in Developments in Time Series Analysis, ed. T. Subba Rao, Chapman and Hall, London. pp. 374-413.

YOUNG, P.C., 2003. Top-down and data-based mechanistic modelling of rainfall-flow dynamics at the catchment scale. Hydrological Processes 17, pp. 2195-2217.

YOUNG, P.C., 2006. Transfer function models, in Encyclopedia of Hydrological Sciences, vol. 3, part II, Rainfall-Runoff Modeling. Anderson, M. (Ed.), John Wiley, Hoboken, N.J. pp. 1985-2000.

YOUNG, P.C., 2011. Recursive Estimation and Time-Series Analysis: An Introduction for the Student and Practitioner. Second Edition. Springer. 505 p.

YOUNG, P.C., BEVEN, K.J., 1994. Data-Based Mechanistic (DBM) modelling and the rainfall-flow nonlinearity. Environmetrics 5 (special issue on 'Environmental Time Series Analysis'), pp. 335-363.

YOUNG, P.C., JAKEMAN, A.J., POST, D.A., 1997. Recent advances in the data-based modelling and analysis of hydrological systems. Water Science and Technology 36 (5), pp. 99-119.

YOUNG, P.C., MCKENNA, P., BRUUN, J., 2001. Identification of non-linear stochastic systems by state dependent parameter estimation. International Journal of Control 74 (18), 1837-1857.

YOUNG, P.C., TAYLOR, C.J., TYCH, W., PEDREGAL, D.J., 2007. The Captain Toolbox. Centre for Research on Environmental Systems and Statistics, Lancaster University, UK. Internet: http://www.lancs.ac.uk/staff/taylorcj/tdc/download. Last accessed: 19.10.2015.

ZADEH, L.A., 1965. Fuzzy sets. Information and Control 8, pp. 338-353.

ZADEH, L.A., 1973. Outline of a new approach to the analysis of complex systems and decision processes. IEEE Transactions on Systems, Man and Cybernetics 1, pp. 28-44.

ZIMMERMAN, H.J., 2001. Fuzzy Set Theory and its Applications, fourth ed. Kluwer Academic Publishers, Boston, MA.

ZOPPOU, C., 2001. Review of urban storm water models. Environmental Modelling & Software 16, pp. 195-231.

Curriculum Vitae

Vincent Wolfs was born on July 25, 1988 in Hasselt, Belgium. He finished his secondary school in Latin-Mathematics at the Virga-Jessecollege in Hasselt in 2006. In June 2009, he received the degree of Bachelor in Civil Engineering from the KU Leuven with the grade magna cum laude. During summer 2010, he did a two-month internship as superintended of maintenance dredging works on the River Paraná in Argentina with the firm Jan De Nul. In June 2011, he received the degree of Master in Civil Engineering from the KU Leuven with the grade magna cum laude. His Master thesis was entitled "Adaptive measures to reduce climate change impacts on flood probabilities", and was nominated for the ie-prize for civil engineering 2011. In September 2011, he enrolled in the PhD program at the Hydraulics Division of the Department of Civil Engineering at the KU Leuven under the supervision of prof. dr. ir. Patrick Willems. In January 2012, he was awarded a 4-year scholarship from the Agency for Innovation by Science and Technology (IWT). During his PhD study, he presented his findings at conferences and workshops in various European countries, China, the United States and Canada. In addition to his PhD research, he was involved in the Civil Engineering Bachelor and Master programs as teaching assistant for the courses Urban Drainage Systems, Open Channel Flow, Hydraulics and Probabilistic Design, and was daily supervisor of 12 Master thesis studies. Since September 2012, he is ombudsperson for the Civil Engineering Bachelor and Master programs. He was also a guest-lecturer of courses organized by VLARIO.

Publications by the author

Articles in internationally reviewed journals

- **Wolfs, V.**, Willems, P. (submitted). Conceptual modelling approach and software for sewer hydraulic simulations. *Advances in Water Resources*.

- **Wolfs, V.**, Meert, P., Willems, P. (2015). Modular conceptual modelling approach and software for river hydraulic simulations. *Environmental Modelling & Software, 71,* 60-77.

- De Vleeschauwer, K., Weustenraad, J., Nolf, C., **Wolfs, V.**, De Meulder, B., Shannon, K., Willems, P. (2014). Green - blue water in the city: quantification of impact of source control versus end-of-pipe solutions on sewer and river floods. *Water Science and Technology, 70* (11), 1825-1837.

- **Wolfs, V.**, Willems, P. (2014). Development of discharge-stage curves affected by hysteresis using time varying models, model tree and neural networks. *Environmental Modelling & Software, 55*, 107-119.

- **Wolfs, V.**, Villazon Gomez, M., Willems, P. (2013). Development of a semi-automated model identification and calibration tool for conceptual modelling of sewer systems. *Water Science and Technology, 68* (1), 167-175.

- **Wolfs, V.**, Willems, P. (2013). A data driven approach using Takagi-Sugeno models for computationally efficient lumped floodplain modeling. *Journal of Hydrology, 503*, 222-232.

Papers at international scientific conferences and symposia, published in full proceedings

- **Wolfs, V.**, Ntegeka, V., Murlà Tuyls, D., Willems, P. (in preparation). Development of computationally efficient urban flood modelling approach. *4th IAHR Europe Congress.* Liege, Belgium, 27-29 July 2016.

- **Wolfs, V.**, Willems, P. (submitted). CMD: fast and tailored conceptual river and sewer models for integrated water management. *9th International Conference Novatech*. Lyon, France, 28 June – 1 July 2016.

- **Wolfs, V.**, Willems, P. (accepted). A flexible and efficient multi-model framework in support of water management. *7th International Water Resources Management Conference of ICWRS*. Bochum, Germany, 18-20 May 2016.

- Vermuyten, E., Meert, P., **Wolfs, V.**, Willems, P. (2015). Using a fast conceptual river model for floodplain inundation forecasting and real-time flood control – a case study in Flanders, Belgium. *21st International Congress on Modelling and Simulation (MODSIM2015).* Broadbeach, Queensland, Australia, 29 November – 4 December 2015.

- **Wolfs, V.**, Willems, P. (2015). Computationally efficient modelling approach and software for sewer quantity simulations. *Proceedings of the 10th conference on urban drainage modelling (10UDM).* International Conference on Urban Drainage Modelling. Québec, Canada, 20-23 September 2015.

- Keupers, I., **Wolfs, V.**, Kroll, S., Willems, P. (2015). Impact analysis of sewer overflows on the receiving river water quality using an integrated conceptual model. *Proceedings of the 10th conference on urban drainage modelling (10UDM).* International Conference on Urban Drainage Modelling. Québec, Canada, 20-23 September 2015.

- **Wolfs, V.**, Willems, P. (2015). Quantification of impact of retention basins on river floods in the Dender catchment in Belgium using computationally efficient models. *36th IAHR World Congress*, Delft/The Hague, The Netherlands, 29 June – 3 July 2015.

- **Wolfs, V.**, Willems, P. (2015). Computationally efficient and flexible modular modelling approach for river and urban drainage systems based

on surrogate conceptual models. EGU General Assembly, Vienna, Austria, 12-17 April 2015; *Geophysical Research Abstracts, 17*, art.nr. EGU2015-9977.

- **Wolfs, V.**, Meert, P., Willems, P. (2014). Semi-automatic conceptual model identification and calibration tool for river hydraulic computations. *River Flow 2014*. River Flow 2014. Lausanne, 3-5 September 2014 (pp. 1-8).

- Vermuyten, E., Van den Zegel, B., **Wolfs, V.**, Meert, P., Willems, P. (2014). Real-time flood control by means of an improved MPC-GA algorithm and a fast conceptual river model for the Demer basin in Belgium. *Proceedings 6th International Conference on Flood Management*. ICFM. Sao Paulo, Brazil, September 2014 (pp. 1-8).

- Van den Zegel, B., Vermuyten, E., **Wolfs, V.**, Meert, P., Willems, P. (2014). Real-time control of floods along the Demer River, Belgium, by means of MPC in combination with GA and a fast conceptual river model. *Proceedings of 11th International Conference on Hydroinformatics (HIC 2014)*. 11th International Conference on Hydroinformatics. New York, USA, 17-21 August 2014 (pp. 1-8).

- **Wolfs, V.**, Willems, P. (2014). Computationally efficient lumped floodplain modelling using data-driven methods. *Proceedings of 11th International Conference on Hydroinformatics (HIC 2014)*. 11th International Conference on Hydroinformatics. New York, USA, 17-21 August 2014 (pp. 1-8).

- Keupers, I., **Wolfs, V.**, Pham, H., Smets, I., Willems, P. (2014). On the role of the WWTP in integrated sewer-WWTP-river impact modelling. *Proceedings 13th International Conference on Urban Drainage*. International Conference on Urban Drainage (13ICUD). Sarawak, Malaysia, 7-12 September 2014 (art.nr. 2518309) (pp. 1-8).

- **Wolfs, V.**, Willems, P. (2013). Development of discharge-stage curves using artificial neural networks and model trees. *Proceedings of the 35th IAHR World Congress*. 35th IAHR World Congress. Chengdu, China, 8-13 September 2013 (art.nr. A10513) (pp. 1-11) Tsinghua University Press.

- De Vleeschauwer, K., Weustenraad, J., **Wolfs, V.**, Willems, P. (2013). Comparison of the up- versus downstream storage on a sewer-river system: a quantitative case study. *Proceedings of the 35th IAHR World*

Congress. 35th IAHR World Congress. Chengdu, China, 8-13 September 2013 (art.nr. A10783) Tsinghua University Press.

- **Wolfs, V.**, Van Steenbergen, N., Willems, P. (2012). Flood probability mapping by means of conceptual modeling. In Muñoz, R. (Ed.), *River Flow 2012: Vol. 2*. International Conference on Fluvial Hydraulics. Costa Rica, 5-7 Sept. 2012 (pp. 1081-1085). London: CRC Press, Taylor & Francis Group.

- **Wolfs, V.**, Villazon Gomez, M., Willems, P. (2012). Development of a semi-automated model identification and calibration tool for conceptual modelling of sewer systems. In Prodanović, D. (Ed.), Plavšić, J. (Ed.), *Proceedings of the 9th conference on urban drainage modelling (9UDM)*. International conference on urban drainage modelling (UDM09). Belgrade, 3-7 September 2012 (pp. 1-11). Belgrade: Faculty of Civil Engineering, University of Belgrade.

Meeting abstracts at international scientific conferences and symposia

- **Wolfs, V.**, Meert, P., Willems, P. (2015). Modular conceptual modelling approach and software for river hydraulic simulations. Workshop on "Reduced complexity modelling". Delft, The Netherlands, 22 May 2015.

- **Wolfs, V.**, Meert, P., Willems, P. (2013). CMD: a semi-automatic conceptual model structure identification and calibration tool for river water quantity modelling. OpenWater Symposium. VUB Brussels, 16-17 Sept. 2013.

Meeting abstracts at national scientific conferences and symposia

- Vissers, S., **Wolfs, V.**, Willems, P. (2011). Adaptieve maatregelen tegen overstromingen in het Denderbekken. ie-prijzen 2011. Antwerpen, 1 December 2011.